Numerical Approximation of Stratonovich SDEs and SPDEs.

Efthalia Tzitzili

Submitted for the degree of
Doctor of Philosophy

Heriot-Watt University

Department of Mathematics,
School of Mathematical and Computer Sciences.

October 2015

Abstract

We consider the numerical approximation of stochastic differential and partial differential equations S(P)DEs, by means of time-differencing schemes which are based on exponential integrator techniques. We focus on the study of two numerical schemes, both appropriate for the simulation of Stratonovich- interpreted S(P)DEs. The first, is a basic strong order 1/2 scheme, called Stratonovich Exponential Integrators (SEI). Motivated by SEI and aiming at benefiting both from the higher order of the standard Milstein scheme and the efficiency of the exponential schemes when dealing with stiff problems, we develop a new Milstein type scheme called Milstein Stratonovich Exponential Integrators (MSEI).

We prove strong convergence of the SEI scheme for high-dimensional semilinear Stratonovich SDEs with multiplicative noise and we use SEI as well as the MSEI scheme to approximate solutions of the stochastic Landau-Lifschitz- Gilbert (LLG) equation in three dimensions. We examine the $L^2(\Omega)$ approximation error of the SEI and MSEI schemes numerically and we prove analytically that MSEI achieves a higher order of convergence than SEI.

We generalise SEI so that it is suited not only for Stratonovich SDEs, but also for Itô and for SDEs interpreted by the 'in-between' calculi. Moreover, we provide a general expression for the predictor contained in SEI and we study the theoretical convergence for the generalised version of the scheme. We show that the order of the scheme used in order to obtain the predictor as well as the stochastic integral interpretation do not affect the overall order of the scheme.

We extend the convergence results for SEI to a space-time context by considering a second order semilinear Stratonovich SPDE with multiplicative noise. We discretise in space with the finite element method and we use SEI for discretising in time. We consider the case where we have trace class noise and we examine analytically the strong order of convergence for SEI.

We implement SEI as a time discretisation scheme and present the results when simulating SPDEs with stochastic travelling wave solutions. Then, we use an alternative method, called 'freezing' method, for approximating wave solutions and estimating the speed of the waves for the stochastic Nagumo and FitzHugh-Nagumo models. The wave position and hence the speed is found by minimising the L^2 distance between a reference function and the travelling wave. While the results obtained from the two different approaches agree, we observe that the behaviour of the wave solution is captured in a smaller computational domain, when we use the freezing method, making it more efficient for long time simulations.

Acknowledgements

This research was supported by the Greek State Scholarships Foundation (IKY) and by a DTA EPSRC grant.

In the first place, I would like to express my special appreciation and thanks to my supervisor Professor Gabriel J. Lord, for his excellent supervision, his valuable guidance and support and above all for being an inspiration to me both as a scientist and as a person.

I would also like to express my gratitude to all of the Mathematics Department faculty members, for teaching me, encouraging me and helping me in various different ways to face the challenges of pursuing a PhD.

I am indebted to all my friends who have supported me throughout my PhD journey and to the occupants of CM G.11 for making my time in Heriot-Watt enjoyable.

Lastly, I would like to thank Konstantinos and Christina for being the most caring parents I could imagine.

Contents

1 Introduction **1**

1.1 Stochastic integral and properties 2

1.2 Numerical schemes for SDEs and convergence 9

1.2.1 Itô Euler-Maruyama scheme 10

1.2.2 Itô Milstein scheme 13

1.2.3 Exponential integrators for Itô SDEs 16

1.2.4 Exponential Milstein-type scheme for Itô SDEs 17

1.2.5 The stochastic Heun method 20

1.2.6 Stratonovich exponential integrators 24

1.3 Numerical comparisons 25

1.4 Some useful tools 27

2 Stratonovich Exponential Integrators **30**

2.1 Strong convergence of SEI in d dimensions 30

2.2 Strong convergence for SEI numerically 42

2.3 Higher order exponential Milstein-type scheme for Stratonovich SDEs 46

2.3.1 MSEI2 46

2.3.2 MSEI 48

3 Generalised version of SEI scheme **63**

3.1 Generalisation of the stochastic integral interpretation 63

3.2 Generalisation of SEI and strong convergence analysis 66

3.3 Numerical experiments on the Heston model 78

4 Stochastic PDEs **81**

4.1 Introduction to space discretization 85

4.1.1 Finite Difference (FD) method 85

4.1.2 Finite Element Method (FEM) 86

4.2 FEM and SEI for reaction-diffusion SPDE 88

5 Stochastic travelling waves **103**
5.1 Simulation of stochastic travelling waves 104
5.2 Freezing method for an SPDE 107
5.2.1 The stochastic Nagumo model 110
5.2.2 The stochastic FitzHugh-Nagumo model 113

6 Conclusion and future work **121**

7 Appendix **123**
7.1 Tensors 123
7.2 Some proofs from Chapters 1 and 3 131
7.3 Matlab codes 135

List of Figures

1.1 Euler-Maruyama approximate solution of the SDE $du = f(u)dt + g(u)dW$, where $f(u) = u$, $g(u) = \frac{1}{2}u$.The solution is approximated in the time interval $[0,1]$ with time step-size 10^{-3} and with initial condition $u_0 = 1$. 12

1.2 Strong convergence rate of Euler-Maruyama scheme with reference time step-size 2^{-14}, $M = 200$ samples. The dashed line is the reference line with slope 1/2 while the solid one shows convergence with order $O(\Delta t^{1/2})$. The plot is on a logarithmic scale. 12

1.3 Reference and approximate time-step comparison of a single sample path. 13

1.4 Approximate solution of $du = \big(Au + f(u)\big)dt + g(u)dW$ with $f = u$, $g(u) = -2u$, $A = 1$, after 10^4 iterations and with initial condition $u_0 = 1$. 17

1.5 Orders of strong convergence for the Itô schemes Euler-Maruyama, Milstein and ETD. $M = 200$ samples, reference step-size 2^{-12}. 17

1.6 Itô exponential Milstein for 3-d Langevin system as given in (2.20). $M = 100$ samples, general i.e. non-diagonal, non-commutative noise. Noise intensity $\sigma = 0.1$, reference step-size 2^{-14}. 20

1.7 Solution approximated by 10^4 iterations of Heun's scheme. $f(u) = u$, $g(u) = 2u$, initial condition $u_0 = 1$. 23

1.8 Approximate solution of a two-dimensional Stratonovich SDE with $\mathbf{F}(\mathbf{u}) = \mathbf{u}$ and $G(\mathbf{u}) = \begin{pmatrix} 2u_1 & 0 \\ 0 & 2u_2 \end{pmatrix}$. The initial condition is $(u_1(0), u_2(0)) = (0.5, 2)$. 24

1.9 Comparison between corrected Euler-Maruyama and stochastic Heun schemes for approximating a Stratonovich SDE with $f(u) = u$, $g(u) = -2u$, $u_0 = 1$. 26

1.10 Comparison between Heun and Stratonovich Exponential Integrators for approximating a Stratonovich SDE with $f(u) = u$, $g(u) = -2u$ and $u_0 = 1$. 26

1.11 Comparison between Euler-Maruyama and ETD for approximating the equivalent Itô SDE of Figures 1.9 and 1.10. 26
1.12 Corrected Euler-Maruyama compared to Exponential Integrators for Stratonovich for approximating an SDE with $f(u) = u$, $g(u) = -2u$ and $u_0 = 1$. 27

2.1 Approximate solution of $du = \big(Au + f(u)\big)dt + g(u) \circ dW$ with $f(u) = u$, $g(u) = -2u$, $A = 1$, after 10^4 iterations and with initial condition $u_0 = 1$. 43
2.2 Green line shows strong convergence of order $O(\sqrt{\Delta t})$ for 2-dimensional system with non-diagonal noise. More specifically, $\mathbf{F}(\mathbf{u}) = 0.2\mathbf{u}$, $G(\mathbf{u}) = \left(\begin{smallmatrix} 0.2u_1 & 2 \\ 2 & 0.2u_2 \end{smallmatrix}\right)$. Bottom blue line shows strong convergence of order $O(\Delta t)$ for 2-dimensional system with diagonal noise. Drift term is the same and diffusion is given by $G(\mathbf{u}) = \left(\begin{smallmatrix} 0.2u_1 & 0 \\ 0 & 0.2u_2 \end{smallmatrix}\right)$. $M = 100$ samples, reference step-size $\frac{T}{N} = 2^{-11}$. 43
2.3 Diagonal noise, three-dimensional example, that is $\mathbf{u} \in \mathbb{R}^3$ and $G(\mathbf{u}) \in \mathbb{R}^{3\times 3}$ is a 3×3 matrix. Strong order of convergence is $O(\Delta t)$ for this example with $M = 300$ samples and $N = 2^{11}$. 44
2.4 LHS:Three-dimensional Langevin system approximated by SEI in blue and by corrected Euler-Murayama in red. RHS: Absolute difference between the two approximations against time, in a log-log scale. Initial condition $\mathbf{u}_0 = (0, 1, 0)^\top$, noise intensity $\sigma = 0.2$, damping parameter $\alpha = 0.25$. 45
2.5 Strong convergence for SEI when solving the three-dimensional Langevin system (2.20). Using polyfit we get that the slope of the fitted line is 0.578590. 46
2.6 strong convergence for SEI solving a three-dimensional system of SDEs similar to (2.20) but with diagonal noise. Diffusion term is given by `ghandle=@(u1,u2,u3) [-sigma*u1 0 0;0 -sigma*u2 0;0 0 -sigma*u3]`. The slope of the fitted line is 1.028944. 46
2.7 Strong convergence for MSEI 2 scheme approximating the LLG system (2.20) with non diagonal, commutative noise against reference line with slope 1. M=100 samples, reference time step-size $\Delta t = 2^{-14}$. The slope of the fitted line is 0.918767. 49
2.8 Comparison between MSEI in black dashed line and Milstein in blue solid line, both for Stratonovich interpreted 3-d Langevin system. Noise intensity $\sigma = 0.1$. 51

2.9 On the left, comparison between MSEI2 (magenta solid line) and MSEI (red dashed line) for approximating the Stratonovich Langevin system. On the right, absolute difference between the two different approximations against time. 52

2.10 Top figures: comparison between MSEI2 (magenta solid line) and corrected exponential Milstein (red solid line), bottom figures: comparison between MSEI (black dashed line) and corrected exponential Milstein (blue solid line), both for approximating Stratonovich Langevin system. Noise intensity $\sigma = 0.1$. 53

2.11 Stratonovich exponential Milstein (MSEI) for 3-d LLG (2.20) with non-diagonal, non-commutative, multiplicative noise without using a predictor. $M = 100$ samples, noise intensity $\sigma = 0.1$, reference time step-size 2^{-14}. The slope of the fitted line is 0.794216. 53

2.12 Left figure shows strong convergence for the MSEI and MSEI2 schemes, both approximating solution to 3-D Langevin system with commutative noise. The slope of the best-fit line for MSEI is 1.075582 while for MSEI2 is 0.982340. Right figure shows the numerical error against computer process time. 54

2.13 Strong convergence for MSEI and MSEI2 schemes on the left, both when approximating solution to 3-D Langevin system with general i.e. non-diagonal, non-commutative noise. Slope of the best-fit line for MSEI is 0.839509 while for MSEI2 is 0.775424. Right figure shows the numerical error against computer process time. 54

2.14 Comparison between the cases where we use a predictor and where we do not use one. The slope of the fitted line for the MSEI predictor case is 0.893925, while for the MSEI no-predictor is 0.812193. 3-d LLG system with non-diagonal, non-commutative noise. Reference time step-size is 2^{-12} and we have considered 500 samples. 55

3.1 Heston stochastic volatility 3/2 model (3.16) approximated by $(1, 1)$-Milstein (3.17) and by general SEI (3.10) for $\alpha = 0.5$. Parameter values are $\mu = 0.1$, $\kappa = 0.2$, $\lambda = \sqrt{0.2}$ and initial condition $u(0) = 0.5$. 79

3.2 Heston stochastic volatility 3/2 model (3.16) approximated by (3.10) for different values $\alpha = 0, 0.5, 1$. Parameter values are $\mu = 0.1$, $\kappa = 0.2$, $\lambda = \sqrt{0.2}$ and initial condition $u(0) = 0.5$. 79

3.3 Heston stochastic volatility 3/2 model (3.16) approximated by (3.10) with $\alpha = 0.5$ and predictors $\tilde{u}_{n+1}$ taken by Euler-Maruyama and ETD schemes (left). Absolute difference between the two different solutions against time (right). Parameter values are $\mu = 0.1$, $\kappa = 0.2$, $\lambda = \sqrt{0.2}$ and initial condition $u(0) = 0.5$. 80

4.1 Order of convergence for SEI and FEM when approximating an SPDE of the form with $f(u) = u(1-u)(u+0.5)$, $g(u) = \sigma u(1-u)$, where $\sigma = 0.5$ is the noise intensity. For (a), reference time step $Dt_{\text{ref}} = 10^{-4}$ with multiples $Dt = \kappa Dt_{\text{ref}}$, $\kappa = (2, 5, 10, 20, 40, 50)$. Fixed spatial step-size 1/63. Slope of the best fitted line, using polyfit, is 0.604180. For (b), reference spatial step $h_{\text{ref}} = 1/(2^9 - 1)$, with multiples $h = \kappa h_{\text{ref}}$, $\kappa = (2, 2^2, 2^3, 2^4, 2^5)$. Fixed time-step size 10^{-4}. For both Figures $M = 100$ samples. 102

5.1 Deterministic Nagumo equation, solved for $t \in [0, 50]$, $x \in [0, 100]$ with Neumann BCs and initial condition $u_0(x) = (1 + e^{-x})^{-1}$. Left: wave solution plot against space at fixed final time. Right: time-space plot of the solution showing the wave travelling with constant speed that can be found by computing the slope of the front. 105

5.2 Deterministic and stochastic fixed time wave solutions plot against space in blue and red respectively. 106

5.3 Stochastic Nagumo equation, solved for $t \in [0, 50]$, $x \in [0, 100]$ with Neumann BCs and initial condition $u_0(x) = (1 + e^{-x})^{-1}$. Left: stochastic wave solution plot against space for fixed time. Right: time-space plot of the solution with the stochastic wave travelling not uniformly in time. 107

5.4 From left to right: (a) Instantaneous speed λ_m against time. (b) Distribution of the instantaneous speed λ_m. (c) Time-averaged speed Λ_m against time. All for one realisation of the stochastic travelling wave solution. 107

5.5 Deterministic Nagumo equation, solved for $t \in [0, 50]$, $x \in [0, 100]$ with Neumann BCs and initial condition $u_0(x) = (1 + e^{-x})^{-1}$. Left: wave-speed λ given directly as an output from (5.13). Right: time-space plot of the solution with the frozen wave. 109

5.6 Stochastic Nagumo equation, solved for $t \in [0, 50]$, $x \in [0, 100]$ with Neumann BCs and initial condition $u_0(x) = (1 + e^{-x})^{-1}$. Left: time-space plot of the travelling wave given by direct simulation of the SPDE. Right: time-space plot of the solution with the frozen wave given by the freezing method which solves the SPDAE. 111

5.7 Given as an output from solving the SPDAE, from left to right: (a) Instantaneous speed λ against time. (b) Distribution of the instantaneous speed λ. (c) Time-averaged speed Λ against time, all for one realisation of the frozen stochastic wave. 111
5.8 Time-averaged speed of the stochastic travelling and frozen wave Λ_m, Λ in blue and green respectively. Noise intensity is 0.1 for the stochastic waves, while the analytic wave speed for the deterministic equation is given in red. 112
5.9 Mean time-averaged speed with error bars. Λ^E for the frozen wave in blue and Λ_m^E for the travelling in green. Mean taken over 100 samples with noise intensity 0.5. 112
5.10 Time-averaged speed at the final computational time against noise intensity. Mean values taken over 1000 different realisations. Speed magnitude increasing as the noise intensity increases. 113
5.11 Travelling wave solution variables u (top) and z (bottom) against space, at fixed times $t^* = 0, T/4, T/2, 3T/4, T$ for $t \in [0, T]$. We solve FHN system (5.18) with noise intensity $\sigma = 0.05$. 114
5.12 Approximate solution of the system (5.18) with noise intensity 0.05. From left to right, solutions u and z against space at fixed time $t^* = T/2$. Time-space plots of the travelling wave solutions u and z. 115
5.13 From left to right: (a) Instantaneous speed λ_m against time. (b) Distribution of the instantaneous speed λ_m. (c) Time-averaged speed Λ_m against time. All for one realisation of the stochastic travelling wave solution. 115
5.14 Mean of the time-averaged speed $\mathbb{E}(\Lambda_m)$ of the travelling wave plot against time. The mean, taken over 20 samples is given by 0.073279. 115
5.15 Freezing method used for approximating FHN model (5.19) with noise intensity $\sigma = 0.1$. From left to right, solutions v and w against space at the final approximating time. Time-space plots of the frozen wave solutions v and w. 117
5.16 From top to bottom, plot of the solutions v and w of the FHN model against space at fixed times $t^* = 0, T/2, T$ for $t \in [0, T]$. Noise intensity $\sigma = 0.1$. 117
5.17 Travelling solution for FHN model (5.18) with noise intensity $\sigma = 0.01$. From left to right, solutions u and z against space at the final approximating time. Time-space plots of the travelling wave solutions u and z. 118

5.18 Frozen solution for FHN model (5.19) with noise intensity $\sigma = 0.01$. From left to right, solutions v and w against space at the final approximating time. Time-space plots of the frozen wave solutions v and w. 118
5.19 Time-averaged speed Λ of the stochastic frozen wave on the left and the deterministic frozen wave on the right. Λ is computed here using the midpoint definition of equation (5.8). 119
5.20 Given as an output of the SPDAE, from left to right for a single realisation: (a) Instantaneous speed of the solution variable $\mathbf{u}$ against time. (b) Distribution of the instantaneous speed of $\mathbf{u}$. (c) Time-averaged speed of $\mathbf{u}$ against time. 119
5.21 Mean of the time-averaged speeds of the frozen wave. The mean, taken over 20 samples, is given by 0.75356. 119

List of Tables

5.1 Instantaneous and time-averaged speeds of both the travelling and the frozen wave cases, for a single realisation of the solution. 111

Chapter 1

Introduction

During the past few decades, there has been an increase in the use of stochastically forced models. As a consequence of the extended use, the complexity of the stochastic differential equations (SDEs) models and sometimes even the absence of exact analytic solutions, the interest for the study of numerical methods that approximate solutions of SDEs has increased as well. Similarly with the deterministic ordinary differential equations, there are numerous factors that have to be considered when developing or choosing an existing scheme to be used for approximating SDEs. These factors among others include the nature of the problem, for example linear, semi-linear, etc, the type of the noise that we have, the size of our system, the accuracy and the time within one wishes it to be solved. This thesis focuses on the development, implementation and analysis of numerical schemes that approximate solutions of SDEs.

We start from this introductory chapter, where we give some background on stochastic calculus and we review the definitions of stochastic integrals interpreted both in Itô and in Stratonovich sense, in multiple dimensions. Next, we set up the context for stochastic differential equations (SDEs), we present the assumptions that determine existence and uniqueness of solutions and we give the connection that exists between Itô and Stratonovich interpreted SDEs.

Driven by the aim to investigate and analyse new numerical schemes for approximating solutions to SDEs, we review some of the standard schemes that exist in the literature and we provide the definitions of two common notions of convergence. Finally, we make a list of inequalities and useful lemmas that we use frequently throughout this thesis.

1.1 Stochastic integral and properties

Let us start from the Itô stochastic integral [34]

$$\int_0^T U(t)dW(t).$$

In the above, $U(t)$ is a stochastic process integrated with respect to a Wiener Process in the interval $[0, T]$. The procedure to be followed for constructing the stochastic integral is similar to the construction of the Riemann integral, in the sense that they are both based on the idea of passing the limit in approximating sums. However, given a partition of the interval in which we integrate, unlike Riemann, stochastic integrals depend on the choice of the intermediate points. As a result, there are several different interpretations, the most widely used of which being the Itô and the Stratonovich integrals. To give the definitions of Itô and Stratonovich integrals and to mention some of their differences, some preliminary definitions are firstly required.

Definition 1.1.1. ***Filtration*** *[54]*

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

i. *A filtration $\{\mathcal{F}_t : t \geqslant 0\}$ is a family of σ-algebras $\subseteq \mathcal{F}$ that are increasing.*

ii. *A filtered probability space is a quadruple $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $\{\mathcal{F}_t : t \geqslant 0\}$ is a filtration of $\mathcal{F}$.*

iii. *If $\{U(t) : t \geqslant 0\}$ is a stochastic process, let $\mathcal{F}_t$ be the smallest σ-algebra such that $U(s)$ is measurable for all $s \leqslant t$. $\{\mathcal{F}_t : t \geqslant 0\}$ is called the natural filtration of $U(t)$.*

The filtration $\mathcal{F}_t$ may be thought of, as all the information provided until time t.

Definition 1.1.2. ***Adapted Processes*** *[14, 39]*

In a filtered probability space $\big(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}\big)$, a real valued stochastic process $\{U(t, \omega) : t \geqslant 0, \omega \in \Omega\}$ is called adapted with respect to $\mathcal{F}_t$, if $U(t, \omega)$ is $\mathcal{F}_t$ measurable for all $t \geqslant 0$.

The Itô integral is initially defined for a class of functions called simple processes. Given a partition $\{t_k\}_{k=0}^{n-1}$ of $[0, T]$, simple processes, also called step processes, take only a finite number of values i.e. $U(t) \equiv U(t_k)$ for $t_k \leqslant t < t_{k+1}$. An additional requirement is that the processes are adapted, hence $U(t_k)$ is $\mathcal{F}_{t_k}-$measurable, for all $k = 0, 1, \ldots, n-1$. Then the Itô integral for simple processes is defined as

$$\int_0^T U(t)dW(t) := \sum_{k=0}^{n-1} U(t_k)\big(W(t_{k+1}) - W(t_k)\big).$$

Next, we extend the above definition to a wider class of processes, called progressively measurable. For $U(t, \omega)$ to be progressively measurable, it means that it is, in a sense, appropriately jointly measurable in the variables t, ω together, see [14, 70] for more. This implies that $U(t, \omega)$ is $\mathcal{F}_t$-measurable and hence $\mathcal{F}_t$-adapted, according to Definition 1.1. Then, the Itô integral is defined in the following way.

Definition 1.1.3. *[70]*
Let $\mathcal{L}^2\big(\Omega, \mathbb{R}\big)$ denote the space of all real valued, progressively measurable processes
$U(t, \omega) : [0, T] \times \Omega \to \mathbb{R}$ such that

$$\mathbb{E}\left[\int_0^T U(t, \omega)^2 dt\right] < \infty.$$

Definition 1.1.4. ***Itô integral in one dimension*** *[63]*
Let $U(t, \omega) \in \mathcal{L}^2(\Omega)$. Then the Itô integral of U is given by

$$\int_0^T U(t, \omega)dW(t) := \lim_{n\to\infty} \int_0^T U_n(t, \omega)dW(t), \;\; \textit{in } L^2(\Omega)$$

where $\{U_n\}_{n\geqslant 0}$ is a sequence of step processes in $\mathcal{L}^2(\Omega)$, such that they converge to U in the mean square sense, i.e.

$$\int_0^T \mathbb{E}\Big[|U(t) - U_n(t)|^2\Big]dt \to 0, \qquad \textit{as } n \to \infty.$$

Such sequence exists for all $U(t, \omega)$ in $\mathcal{L}^2(\Omega)$, for details see [63]. The one-dimensional definition can now be extended to higher dimensions in the following way.

Definition 1.1.5. ***Itô integral in high dimensions***
Let $\boldsymbol{W} = (W_1, W_2, \ldots, W_m)$ be an $m-$dimensional Wiener process and let $\mathcal{L}^2_{\mathcal{H}} = \mathcal{L}^2(\Omega, \mathbb{R}^{d\times m})$ be the set of $\mathbb{R}^{d\times m}$-processes $u = [u_{ij}(t, \omega)]$ such that u_{ij} belongs to $\mathcal{L}^2(\Omega, \mathbb{R})$ for all $i = 1, \ldots, d$, $j = 1, \ldots, m$. Then the Itô integral in n dimensions is given by

$$\int_0^T u(t, \omega)d\boldsymbol{W}(t) := \lim_{n\to\infty} \int_0^T u_n(t, \omega)d\boldsymbol{W}(t), \;\; \textit{in } L^2(\Omega).$$

Where

$$\int_0^T u(t,\omega)d\boldsymbol{W}(t) = \int_0^T \begin{pmatrix} u_{11} & \dots & u_{1m} \\ \vdots & \ddots & \vdots \\ u_{d1} & \dots & u_{dm} \end{pmatrix} \begin{pmatrix} dW_1 \\ \vdots \\ dW_m \end{pmatrix}$$

is a column vector in $L^2(\Omega, \mathbb{R}^d)$ *with* $i-$*th component given by*

$$\sum_{j=1}^{m} \int_0^T u_{ij}(t)dW_j(t).$$

Two important properties of the Itô integral are the Itô's isometry and the martingale property as seen below in Proposition 1.1.6. For a proof of the proposition see [63].

Proposition 1.1.6. ***Itô integral properties*** *[63]*

Let $u(t,\omega) \in \mathcal{L}^2_{\mathcal{H}}$, *then the Itô integral satisfies the following properties.*

1. *Itô's isometry*

$$\mathbb{E}\Big[\| \int_0^T u(t)d\boldsymbol{W}(t)\|_2^2\Big] = \int_0^T \mathbb{E}\big[\|u(t)\|_F^2\big]dt.$$

2. *Martingale Property*

$$\mathbb{E}\left[\int_0^T u(t)d\boldsymbol{W}(t)\right] = 0.$$

Note that when we deal with matrices we use the standard Frobenius norm for matrices, defined by summing up the absolute squares of the elements of the matrix and then taking the square root of the sum. We denote this by $\|\cdot\|_F$.

Having defined the Itô interpretation of the stochastic integral, let us now see how the stochastic integral is defined in the Stratonovich case. Similarly with the Itô case, we consider a partition $\{t_k\}_{k=0}^n = \{0 = t_0 < t_1 < \ldots < t_n = T\}$ of $[0,T]$. We now evaluate the integrand $U(t)$ at the midpoint $\frac{t_k + t_{k+1}}{2}$ of each subinterval $[t_k, t_{k+1}]$. That is we approximate $U(t)$ by $\frac{U(t_{k+1})+U(t_k)}{2}$. We also consider the so called Brownian increments $\Delta W_k = W(t_{k+1}) - W(t_k)$, and we define

$$Z_n := \sum_{k=0}^{n-1} \frac{U(t_{k+1}) + U(t_k)}{2} \Delta W_k.$$

Then, the Stratonovich integral of $U(t)$ with respect to the Wiener process $W(t)$, is taken by the following limit

$$\int_0^T U(t) \circ dW(t) = \lim_{n\to\infty} Z_n.$$

Note that, throughout this thesis we use the above notation and more specifically the $\circ$ symbol, to distinguish the Stratonovich from the Itô integral. So, the definition of the Stratonovich integral is given below.

Definition 1.1.7. ***Stratonovich integral in one dimension*** *[78, 42, 4]*
Let $W(t)$ be a Wiener process and $\varphi : \mathbb{R} \times [0,T] \to \mathbb{R}$ is a C^2 function such that

$$\mathbb{E}\Big[\int_0^T |\varphi(W,t)|^2 dt\Big] < \infty,$$

then

$$\int_0^T \varphi(W,t) \circ dW := \lim_{n\to\infty} \sum_{k=0}^{n-1} \varphi\left(\frac{W(t_{k+1}) + W(t_k)}{2}, t_k\right)\big(W(t_{k+1}) - W(t_k)\big).$$

Alternatively, one can define the Stratonovich integral using as a reference the Itô one. We give this approach in the following definition which is also extended in higher dimensions.

Definition 1.1.8. ***Stratonovich integral in high dimensions***
Let $\boldsymbol{W}(t)$ be an m-dimensional Wiener process and let $\varphi : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d\times m}$ be a C^2 function such that

$$\mathbb{E}\Big[\int_0^T |\varphi(\boldsymbol{u},t)|^2 dt\Big] < \infty,$$

where $\boldsymbol{u}(t)$ is a real valued stochastic process in $L^2(\Omega, \mathbb{R}^d)$. Then, the Stratonovich integral is defined by

$$\int_0^T \varphi(\boldsymbol{u}(t)) \circ d\boldsymbol{W}(t) := \int_0^T \varphi(\boldsymbol{u}(t)) d\boldsymbol{W}(t) + \frac{1}{2}[\varphi(\boldsymbol{u}), \boldsymbol{W}].$$

Where $[\varphi(\boldsymbol{u}), \boldsymbol{W}]$ is the quadratic variation, given by the following limit in $L^2(\Omega)$

$$[\varphi(\boldsymbol{u}), \boldsymbol{W}] := \lim_{\Delta t\to 0} \sum_{k=0}^{n-1} \big(\varphi(\boldsymbol{u}(t_{k+1})) - \varphi(\boldsymbol{u}(t_k))\big)\big(\boldsymbol{W}(t_{k+1}) - \boldsymbol{W}(t_k)\big),$$

with $\Delta t := \max_k (t_{k+1} - t_k)$.

Definition 1.1.8 implies that there is a connection between the two different interpretations of the stochastic integral. This connection is given by the so called Itô-Stratonovich correction that we see later on in this chapter.

Apart from the connection though, there are some differences between the two integrals. The more obvious difference is the choice of the intermediate points. As seen in the construction process of the integral, the integrand is evaluated at the left-end point in case of the Itô integral while it is evaluated at the midpoint in case of Stratonovich. In fact, it can be evaluated at any chosen point and this choice determines a new integral. We generalise the definition of stochastic integral in Chapter 3. Another difference is that, unlike Itô, the Stratonovich integral has some properties which are similar to those of the classical calculus. More specifically, the Stratonovich integral is constructed in a way so that the standard-calculus chain rule holds. This property makes the Stratonovich integral a natural choice for a variety of models in physics and computational biology. Some examples of such models can be found in [16, 42]. On the other hand, the Stratonovich integral does not have the martingale property. This, makes the Itô integral widely used in stochastic financial analysis.

After reviewing the definitions of the Itô and Stratonovich integrals, we can now consider the one-dimensional Itô SDE given in the following form

$$du = f(u)dt + g(u)dW. \tag{1.1}$$

Here $u(t,\omega)$ is a real-valued stochastic process in $L^2(\Omega,\mathbb{R})$, note that for simplicity's sake we often use the notation $u(t)$ or u instead of $u(t,\omega)$. The functions f and g both take values in $\mathbb{R}$ for the one-dimensional case and are often called drift and diffusion terms respectively.If $g(u)$ is a constant, we say that we have additive noise whereas if $g(u)$ depends on u we say that we have multiplicative noise. The same applies to the high dimension Itô SDE

$$d\mathbf{u} = \mathbf{F}(\mathbf{u})dt + G(\mathbf{u})d\mathbf{W}, \tag{1.2}$$

where $\mathbf{u} = (u_1,\dots,u_n) \in \mathbb{R}^d$, $\mathbf{F} : \mathbb{R}^d \to \mathbb{R}^d$ and $G : \mathbb{R}^d \to \mathbb{R}^{d\times m}$, that is

$$G(u_1,\dots,u_n) = \begin{pmatrix} G_{11}(\mathbf{u}) & \dots & G_{1m}(\mathbf{u}) \\ \vdots & \ddots & \vdots \\ G_{d1}(\mathbf{u}) & \dots & G_{dm}(\mathbf{u}) \end{pmatrix}.$$

Also, $\mathbf{W}(t) = (W_1(t),\dots,W_m(t))$ with $W_k(t)$ being independent of $W_\ell(t)$ for $k \neq \ell$ and $k,\ell = 1,\dots,m$. Next, we establish the context in which we are working, by including a list of assumptions that, among others, ensure the existence and uniqueness of a solution for the type of SDEs that we consider.

Assumption 1.1.9. ***Lipschitz condition***

The functions $\boldsymbol{F} : \mathbb{R}^d \to \mathbb{R}^d$, $G : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ *satisfy the Lipschitz condition. That is, there exist constants* $K, L > 0$ *such that for* $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d$

$$\|\boldsymbol{F}(\boldsymbol{u}) - \boldsymbol{F}(\boldsymbol{v})\|_2^2 \leqslant K\|\boldsymbol{u} - \boldsymbol{v}\|_2^2$$

and

$$\|G(\boldsymbol{u}) - G(\boldsymbol{v})\|_F^2 \leqslant L\|\boldsymbol{u} - \boldsymbol{v}\|_2^2.$$

Assumption 1.1.10. ***Linear growth condition***

Both $\boldsymbol{F}$ *and* G *satisfy the linear growth condition. That is there exist constants* $K, L > 0$ *such that for* $\boldsymbol{u} \in \mathbb{R}^d$

$$\|\boldsymbol{F}(\boldsymbol{u})\|_2^2 \leqslant K\big(1 + \|\boldsymbol{u}\|_2^2\big)$$

and

$$\|G(\boldsymbol{u})\|_F^2 \leqslant L\big(1 + \|\boldsymbol{u}\|_2^2\big).$$

Then, under Assumptions 1.1.9 and 1.1.10 there exists a unique solution to (1.2) that satisfies the following integral equation

$$\mathbf{u}(t) = \mathbf{u}(0) + \int_0^t \mathbf{F}(\mathbf{u}(s))ds + \int_0^t G(\mathbf{u}(s))d\mathbf{W}(s) \tag{1.3}$$

or for $d = m = 1$ the solution u satisfies

$$u(t) = u(0) + \int_0^t f\big(u(s)\big)ds + \int_0^t g\big(u(s)\big)dW(s). \tag{1.4}$$

For a proof of the existence and uniqueness we refer to [42, 54, 58, 63]. Note that the Lipschitz and the linear growth conditions are strong assumptions to impose on the drift and diffusion terms that restrict the class of the problems for which we can prove existence and uniqueness. Therefore, there is a big interest in finding ways to 'relax' these assumptions. There is not any work towards this direction included in this thesis, however, some approaches can be found in [24, 59, 80].

Now, we look back at Definition 1.1.8 where apart from the Stratonovich integral, we also review the definition of the quadratic variation. Under the assumptions that $\varphi(u)$ is a C^2 function and that u is a solution of equation (1.1), it turns out that for $d = m = 1$

$$[\varphi(u), W] = \int_0^t \varphi'(u(s))g(u(s))ds.$$

Then,

$$\int_0^t \varphi(u(s)) \circ dW(s) = \int_0^t \varphi(u(s))dW(s) + \frac{1}{2}\int_0^t \varphi'(u(s))g(u(s))ds.$$

In particular for $\varphi(u) = g(u)$, we have that

$$\int_0^t g(u(s)) \circ dW(s) = \int_0^t g(u(s))dW(s) + \frac{1}{2}\int_0^t g'(u(s))g(u(s))ds. \tag{1.5}$$

On the other hand, (1.4) can be equivalently written as

$$u(t) = u(0) + \int_0^t \tilde{f}(u(s))ds + \int_0^t g(u(s))dW(s) + \frac{1}{2}\int_0^t g'(u(s))g(u(s))ds, \tag{1.6}$$

where $\tilde{f}(u(s)) = f(u(s)) - \frac{1}{2}g'(u(s))g(u(s))$. Then, by (1.5) we have that (1.6), and hence (1.4) is equivalent to

$$u(t) = u(0) + \int_0^t \tilde{f}(u(s))ds + \int_0^t g(u(s)) \circ dW(s). \tag{1.7}$$

Note that, equation (1.7) is the integral form of the one-dimensional Stratonovich SDE

$$du = \tilde{f}(u)dt + g(u) \circ dW. \tag{1.8}$$

This connection between (1.4) and (1.7) enables the user to convert an Itô SDE to its equivalent Stratonovich SDE and vice versa. We call it Itô-Stratonovich correction because the transformation is achieved simply by switching or 'correcting' the drift term from f to $\tilde{f}$ or from $\tilde{f}$ to f in the opposite case. For more details we refer to [54].

The Itô-Stratonovich correction can be extended to higher dimensions as well. More specifically, the i-th component of the Itô SDE

$$du_i = F_i(\mathbf{u})dt + \sum_{j=1}^{m} G_{ij}(\mathbf{u})dW_j(t), \;\; i = 1, \dots, d$$

with integral form

$$u_i(t) = u_i(0) + \int_0^t F_i(\mathbf{u}(s))ds + \sum_{j=1}^{m}\int_0^t G_{ij}(\mathbf{u}(s))dW_j$$

is equivalent to the corresponding i-th component of the Stratonovich equation

$$du_i = \tilde{F}_i(\mathbf{u})dt + \sum_{j=1}^{m} G_{ij}(\mathbf{u}) \circ dW_j(t)$$

with integral form

$$u_i(t) = u_i(0) + \int_0^t \tilde{F}_i(\mathbf{u}(s))ds + \sum_{j=1}^{m} \int_0^t G_{ij}(\mathbf{u}(s)) \circ dW_j(s), \tag{1.9}$$

where $\tilde{F}_i(\mathbf{u}(s)) = F_i(\mathbf{u}(s)) - \frac{1}{2}\sum_{j=1}^{m} \frac{\partial G_{ij}(\mathbf{u})}{\partial u_i}$.

Note that existence and uniqueness of solution to the Stratonovich SDE are satisfied by imposing Assumptions 1.1.9 and 1.1.10 to $\tilde{F}_i$, as we did to F_i in the Itô case. This implies that further to $F_i(\mathbf{u})$ we also require that $G(\mathbf{u})$ is differentiable with Lipschitz derivative that satisfies linear growth. For more details, see for example [45] where the authors prove existence of a unique solution in the case where $G(\mathbf{u})$ satisfies the Frobenius condition or [50] where existence and uniqueness of a Stratonovich-interpreted SDE is proved for the case of scalar processes.

From now on when we work with Stratonovich SDEs we use the following d-dimensional equation

$$d\mathbf{u} = \mathbf{F}(\mathbf{u})dt + G(\mathbf{u}) \circ d\mathbf{W} \tag{1.10}$$

and we require that $G(\mathbf{u})$ is differentiable with Lipschitz derivative that also satisfies the linear growth condition. Then there exists a unique solution that satisfies

$$\mathbf{u}(t) = \mathbf{u}(0) + \int_0^t \mathbf{F}(\mathbf{u}(s))ds + \int_0^t G(\mathbf{u}(s)) \circ d\mathbf{W}(s). \tag{1.11}$$

1.2 Numerical schemes for SDEs and convergence

In this section, we present a selection of standard numerical schemes for SDEs together with some of their features. What we call a scheme, basically consists of a time discretization technique which is applied to the continuous form of the SDE that we are dealing with. Towards this direction, we firstly consider a discretisation of the time interval $[0, T]$ by defining a step-size $\Delta t = \frac{T}{N}$, with N being an integer standing for the number of steps. Then, we look at the solution of the SDE at times $t_n = n\Delta t$ and we denote the exact solution at time t_n by $u(t_n)$, while the approximate solution, given by some numerical scheme, will be denoted as u_n.

Once a scheme has been applied, the more natural question to ask is how efficient it is, that is how well and how fast it approximates the exact solution, assuming the latter exists. Before introducing any numerical schemes for SDEs, we need to define a notion of their convergence. Throughout this thesis, we are using the concept of strong convergence which concerns finding how big the mean of the difference between the exact and the approximate solution can become. When we deal with a numerical scheme that converges strongly to our exact solution, we expect to get that the mean of the difference described above, goes to zero as Δt goes to zero. The order of convergence is given in the following definition, for more details we refer to [21, 22, 43, 54].

Definition 1.2.1. ***Strong Convergence*** *[43]*

Let $\boldsymbol{u}_n$ be the numerical approximation of the exact solution $\boldsymbol{u}(t_n)$ of an SDE considered in the time interval $[0,T]$. A method is said to have strong convergence equal to γ if there exists a constant C, independent of Δt, such that

$$\sup_{0\leqslant t_n\leqslant T} \left\| \boldsymbol{u}(t_n) - \boldsymbol{u}_n \right\|_{L^2(\Omega,\mathbb{R}^d)} \leq C\Delta t^{\gamma} \tag{1.12}$$

for Δt sufficiently small, $T = n\Delta t$ and with $\| \cdot \|_{L^2(\Omega,\mathbb{R}^d)} := \left(\mathbb{E}\big[\| \cdot \|_2^2\big]\right)^{1/2}$

Note that in the one-dimensional case, the above definition reduces to the mean square absolute difference, i.e.

$$\sup_{0\leqslant t_n\leqslant T} \mathbb{E}\big[|u(t_n) - u_n|\big] \leq C\Delta t^{\gamma}.$$

Another notion of convergence is the weak convergence that concerns the average solution behaviour. In order to examine weak convergence, one has to firstly choose an appropriate function Φ that must satisfy some polynomial smoothness conditions and then look at the weak error which measures the difference between the mean of $\Phi(u_n)$ and the mean of $\Phi(u(t_n))$. This essentially means that weak convergence measures in a sense the error of the means in contrast to strong convergence that measures the mean of the error. Throughout this thesis we do not examine weak convergence of the presented numerical schemes, however, for completeness purposes we provide the following definition and we refer to [21, 22, 42, 43] for further reading.

Definition 1.2.2. ***Weak Convergence*** *[43]*

Let u_n be the numerical approximation of the exact solution $u(t_n)$ of an SDE considered in the time interval $[0,T]$. A method is said to have weak convergence equal to γ if there exists a constant C such that for all functions Φ in some class

$$|\mathbb{E}(\Phi(u_n)) - \mathbb{E}(\Phi(u(T)))| \leq C\Delta t^{\gamma} \tag{1.13}$$

for Δt sufficiently small and $T = n\Delta t$.

1.2.1 Itô Euler-Maruyama scheme

One of the most standard methods for numerically approximating solutions of Itô stochastic differential equations is the Euler-Maruyama method. The Euler-Maruyama or stochastic Euler method is an altered version of the standard deterministic Euler method so that it can be applied to SDEs.

Consider the Itô SDE (1.2) and let Assumptions 1.1.9 and 1.1.10 hold. Then, the solution of (1.2) considered in the discretised time intervals $[t_n, t_{n+1}]$ satisfies the following integral equation

$$\mathbf{u}(t_{n+1}) = \mathbf{u}(t_n) + \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{u}(s))ds + \int_{t_n}^{t_{n+1}} G(\mathbf{u}(s))d\mathbf{W}(s). \tag{1.14}$$

In order to derive the Euler-Maruyama scheme, we approximate the functions $\mathbf{F}, G$ in the intervals $[t_n, t_{n+1}]$, by taking their Taylor expansions around $\mathbf{u}(t_n)$, see more in [42]. Recall from the definition of the Itô stochastic integral that we evaluate the integrant at the left-end point of each subinterval and note that our choice of $\mathbf{u}(t_n)$ here aims at being consistent with this definition. This observation is important for understanding the difference later when we derive schemes for Stratonovich SDEs. We denote the time-step $(t_{n+1} - t_n)$ by Δt and the Brownian increments $\big(\mathbf{W}(t_{n+1}) - \mathbf{W}(t_n)\big)$ by $\Delta \mathbf{W}_n$. Hence the Euler-Maruyama scheme is given by

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \mathbf{F}(\mathbf{u}_n)\Delta t + G(\mathbf{u}_n)\Delta \mathbf{W}_n. \tag{1.15}$$

An example of using the Euler-Maruyama scheme for approximating the solution of an SDE can be seen in Figure (1.1). It turns out that the Euler-Maruyama scheme converges with a strong order $\mathcal{O}(\Delta t^{1/2})$, which can be seen in Figure 1.2. We should not fail to mention that Assumptions 1.1.9 and 1.1.10 are essential for the strong convergence of Euler-Maruyama. Should these not be satisfied, strong convergence is not guaranteed, see [32].

The idea behind examining strong convergence numerically, is to consider a small Δt_{ref} as a reference time step and then use this to create bigger increments which are multiples of the reference one. Next, we measure the difference between the approximate solution when Δt is relatively big and the solution when we use Δt_{ref}. We repeat this procedure several times and we take the mean of all the differences recorded. This gives us an estimate of the error between the exact and the approximate solution. In Figure 1.3 we can see the difference between the reference and the approximate solution of one sample path. The Algorithms that have been used for generating Figures 1.1 and 1.2 can be found in the corresponding section of the Appendix.

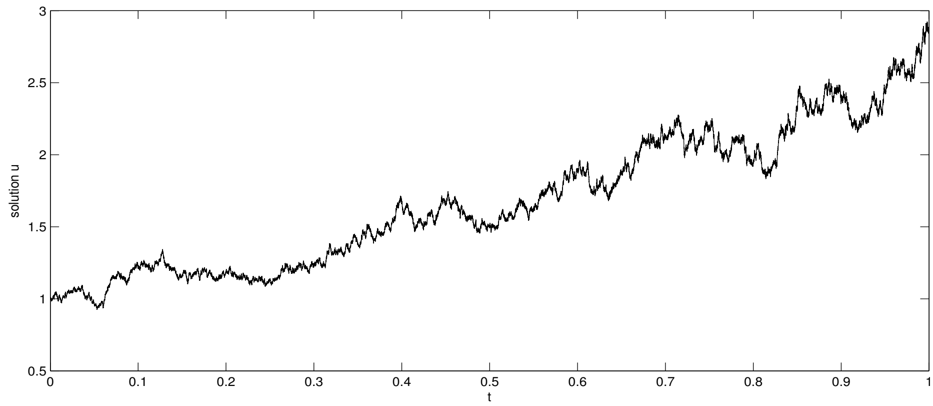


Figure 1.1: Euler-Maruyama approximate solution of the SDE $du = f(u)dt + g(u)dW$, where $f(u) = u$, $g(u) = \frac{1}{2}u$.The solution is approximated in the time interval $[0, 1]$ with time step-size 10^{-3} and with initial condition $u_0 = 1$.

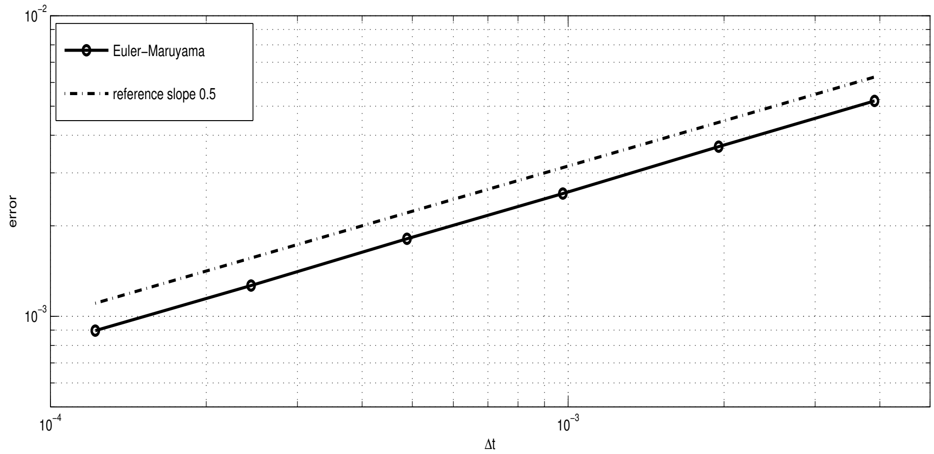


Figure 1.2: Strong convergence rate of Euler-Maruyama scheme with reference time step-size 2^{-14}, $M = 200$ samples. The dashed line is the reference line with slope 1/2 while the solid one shows convergence with order $O(\Delta t^{1/2})$. The plot is on a logarithmic scale.

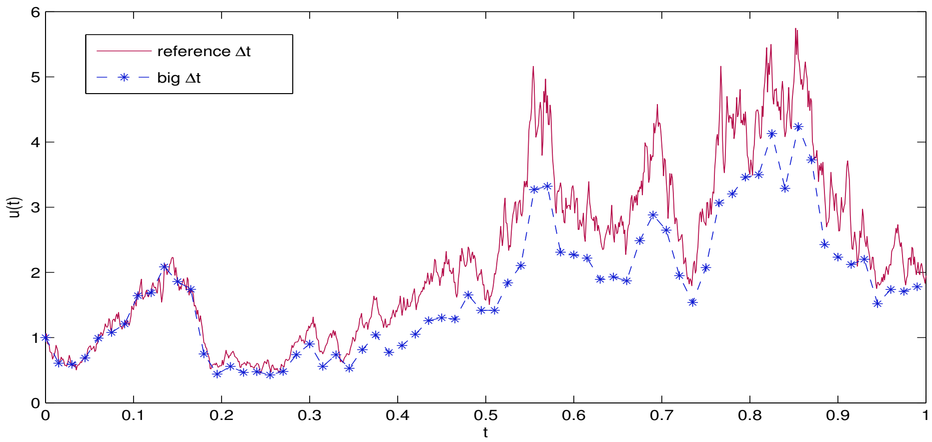


Figure 1.3: Solution approximated by the Euler-Murayama scheme for two different sizes of Δt. The red path is considered as the reference path while the blue one as the approximate. See Appendix, algorithm 7.3.2 for $M = 1$.

1.2.2 Itô Milstein scheme

Another standard higher order method is Milstein's method. We start from looking at the solution of equation (1.2) considered in the time interval $[r, t]$

$$\mathbf{u}(t) = \mathbf{u}(r) + \int_r^t \mathbf{F}(\mathbf{u}(s))ds + \int_r^t G(\mathbf{u}(s))d\mathbf{W}(s).$$

Similarly with Euler-Maruyama, we evaluate $\mathbf{F}(\mathbf{u}(s))$ at the left-end point $\mathbf{u}(r)$ and we Taylor expand $G(\mathbf{u}(s))$ about $\mathbf{u}(r)$. Note that, for the diffusion term $G(\mathbf{u}(s))$, we now use the higher order approximation

$$G(\mathbf{u}(s)) = G(\mathbf{u}(r)) + DG(\mathbf{u}(r))\big(\mathbf{u}(s) - \mathbf{u}(r)\big) + R_T,$$

where R_T is the Taylor expansion remainder term and $DG(\mathbf{u}) : \mathbb{R}^{d\times d} \to \mathbb{R}^{d^3}$, for $d = m$, is a rank-three tensor given by

$$DG(\mathbf{u}) = \begin{pmatrix} \nabla G_{11}(\mathbf{u}) & \dots & \nabla G_{1d}(\mathbf{u}) \\ \vdots & \ddots & \vdots \\ \nabla G_{d1}(\mathbf{u}) & \dots & \nabla G_{dd}(\mathbf{u}) \end{pmatrix}.$$

Note that $\big(\mathbf{u}(s) - \mathbf{u}(r)\big) \in \mathbb{R}^d$, so the product $DG(\mathbf{u}(r))\big(\mathbf{u}(s) - \mathbf{u}(r)\big)$ belongs to $\mathbb{R}^{d\times d}$. More details of this operation and on tensors' background can be found in the corresponding section of the Appendix. Then, using the Euler-Maruyama scheme, we substitute for $\big(\mathbf{u}(s) - \mathbf{u}(r)\big)$ to obtain

$$
\begin{aligned}
\mathbf{u}(t) = \mathbf{u}(r) &+ \int_r^t \mathbf{F}(\mathbf{u}(r))ds + \int_r^t G(\mathbf{u}(r))d\mathbf{W}(s) \\
&+ \int_r^t DG(\mathbf{u}(r))\Big(G(\mathbf{u}(r)) \int_r^s d\mathbf{W}(\tau)\Big)d\mathbf{W}(s) + R_M,
\end{aligned} \tag{1.16}
$$

where R_M is a remainder term given by

$$
R_M = R_T + R_E + \int_r^t DG(\mathbf{u}(r))\Big(\mathbf{F}(\mathbf{u}(r)) \int_r^s d\tau\Big)d\mathbf{W}(s).
$$

Then, equation (1.16) considered in the time intervals $[t_n, t_{n+1}]$ reads

$$
\begin{aligned}
\mathbf{u}(t_{n+1}) = \mathbf{u}(t_n) &+ \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{u}(t_n))ds + \int_{t_n}^{t_{n+1}} G(\mathbf{u}(t_n))d\mathbf{W}(s) \\
&+ \int_{t_n}^{t_{n+1}} DG(\mathbf{u}(t_n))\Big(G(\mathbf{u}(t_n)) \int_{t_n}^s d\mathbf{W}(\tau)\Big)d\mathbf{W}(s) + R_M.
\end{aligned} \tag{1.17}
$$

Then, the double integral can be expanded according to Example 3 of the Appendix and the i-th component of the Itô Milstein scheme is given by

$$
\begin{aligned}
u_{i,n+1} = u_{i,n} + F_i(\mathbf{u}_n) + \sum_{\ell=1}^{d} G_{i\ell}(\mathbf{u}_n)\Delta W_{\ell,n} + \frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{d} \frac{\partial G_{ij}}{\partial u_k}(\mathbf{u}_n)G_{kj}(\mathbf{u}_n)\big(\Delta W_{j,n}^2 - \Delta t\big) \\
+ \frac{1}{2}\sum_{j<\ell=1}^{d}\sum_{k=1}^{d}\Big(\frac{\partial G_{i\ell}}{\partial u_k}(\mathbf{u}_n)G_{kj}(\mathbf{u}_n) + \frac{\partial G_{ij}}{\partial u_k}(\mathbf{u}_n)G_{k\ell}(\mathbf{u}_n)\Big)\Delta W_{j,n}\Delta W_{\ell,n} \\
+ \frac{1}{2}\sum_{j<\ell=1}^{d}\sum_{k=1}^{d}\Big(\frac{\partial G_{i\ell}}{\partial u_k}(\mathbf{u}_n)G_{kj}(\mathbf{u}_n) + \frac{\partial G_{ij}}{\partial u_k}(\mathbf{u}_n)G_{k\ell}(\mathbf{u}_n)\Big)A_{j\ell,n},
\end{aligned} \tag{1.18}
$$

where $A_{j\ell,n}$ is called the Lévy area of (W_j, W_ℓ) and is given in general by the following expression

$$
A_{ij,n} = A_{ij}(t_n, t_{n+1}) = \int_{t_n}^{t_{n+1}}\int_{t_n}^{s} dW_i(\tau)dW_j(s) - \int_{t_n}^{t_{n+1}}\int_{t_n}^{s} dW_j(\tau)dW_i(s). \tag{1.19}
$$

This, represents the area enclosed between the curve of $t \in [t_n, t_{n+1}] \to (W_i(t), W_j(t))$ and its chord, see [53].

In order to implement Milstein's scheme numerically, one has to simulate the Lévy area, that is the double stochastic integrals of (1.19). There are several ways for simulating double or in general multiple stochastic integrals in the literature. One of these methods is based on a Karhunen-Loève or Fourier series expansion of a Brownian bridge process. For this method for the Itô case we refer to [44] while for Stratonovich SDEs we refer to [42] where the authors use the Fourier expansion to obtain an approximation of a Brownian bridge process and then they show that integrals with respect to such a process converge to Stratonovich integrals. An application of the same method for the approximation of Itô multiple integrals, that arise in a chemical Langevin system, can be found in [33].

Another way to define multiple stochastic integrals against Brownian motion is the use of the theory of rough paths. In order to do this, one has to construct an approximation that converges to a given Brownian trajectory. It is not within the scope of this thesis though to further examine this method, for more details we refer to [52, 53] and references therein.

In our approach, we simulate the double stochastic integrals that arise when dealing with higher order schemes, according to a method suggested by [40] and implemented in [23]. The idea behind this method, is to use a simpler numerical scheme for simulating the integrals themselves. In particular, let $I_{i,j}(t_n, t_{n+1})$ denote the double integral $\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW_i(\tau) dW_j(s)$. Then, $I_{2,1}(t_n, t_{n+1})$ for example, denotes the integral $\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW_2(\tau) dW_1(s)$. We consider the two-dimensional SDEs system

$$\begin{aligned} du_1 &= u_2 dW_1 \\ du_2 &= dW_2, \end{aligned} \tag{1.20}$$

over the time interval $[t_n, t_{n+1}]$ with initial condition $(0, W_2(t_n))$. It is easy to see that the solution of (1.20) at time $t = t_{n+1}$ gives $I_{2,1}(t_n, t_{n+1})$, that is $u_1(t_{n+1}) = I_{2,1}(t_n, t_{n+1})$, $u_2(t_{n+1}) = \Delta W_{2,n} = W_2(t_{n+1}) - W_2(t_n)$. So, we can use for example the Euler-Maruyama scheme for solving the system (1.20) and therefore obtain an approximation for the stochastic integral. The Euler-Maruyama approximate solution in our case is given by

$$\begin{aligned} u_{1,n+1} &= u_{1,n} + u_{2,n} \Delta W_{1,n} \\ u_{2,n+1} &= u_{2,n} + \Delta W_{2,n}. \end{aligned} \tag{1.21}$$

We rename $u_{1,n+1}$ to $v_{1,n+1}$ to conclude that the double integral term in Milstein scheme is given by

$$A_{12,n} = I_{12}(t_n, t_{n+1}) - I_{21}(t_n, t_{n+1}) = \hat{v}_{1,n+1} - v_{1,n+1},$$

where $\hat{v}_{1,n+1}$ is obtained in a similar way with $v_{1,n+1}$ just by interchanging the indices 1 and 2 in (1.20).

Milstein's method converges with a strong order of $O(\Delta t)$, see Figure 1.5. When we simulate the double stochastic integrals by means of the auxiliary system above, the strong order of convergence remains $O(\Delta t)$, under the assumption that we use the same step-size both for the Euler-Maruyama approximation of (1.20) and for the numerical scheme (1.18).

1.2.3 Exponential integrators for Itô SDEs

The next scheme is an extension of the standard deterministic exponential time differencing schemes. These schemes, also found in the literature as exponential integrators, are known to be efficient for solving large systems of stiff differential equations, see for example [26] and [12]. This, is achieved by bringing the original problem to a semilinear form and combining the exact solution of the linear part with an explicit approximation of the non-linear one. For more information related to deterministic exponential schemes we refer to [27, 28, 29]. The extension to SDEs was considered in [36] while the extension to SPDEs was considered in [55].

We consider a semilinear Itô SDE of the form (1.2), with drift given by $A\mathbf{u}+\mathbf{F}(\mathbf{u})$

$$d\mathbf{u} = \big(A\mathbf{u} + \mathbf{F}(\mathbf{u})\big)dt + G(\mathbf{u})d\mathbf{W}, \tag{1.22}$$

where $\mathbf{F}(\mathbf{u})$ is in general nonlinear and satisfies Assumptions 1.1.9 and 1.1.10 as before and $A \in \mathbb{R}^{d\times d}$ is a linear bounded operator.

The Itô exponential integrators scheme that we review in this section is called SETD and can be derived by using the variation of constants formula, as applied in [1]. Then, a solution of (1.22), considered in the time interval $[t_n, t_{n+1}]$ satisfies

$$\mathbf{u}(t_{n+1}) = e^{\Delta t A}\mathbf{u}(t_n) + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{F}(\mathbf{u}(s))ds + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}G(\mathbf{u}(s))d\mathbf{W}(s). \tag{1.23}$$

We evaluate $\mathbf{F}(\mathbf{u}(s))$ and $G(\mathbf{u}(s))$ at the left-end point $\mathbf{u}(t_n)$ and we approximate them by $\mathbf{F}(\mathbf{u}(t_n))$ and $G(\mathbf{u}(t_n))$ respectively. We compute the deterministic integral and approximate the stochastic one so that one version of the SETD scheme is given by

$$\mathbf{u}_{n+1} = e^{\Delta t A}\mathbf{u}_n + \phi(\Delta t A)\mathbf{F}(\mathbf{u}_n) + e^{\Delta t A}G(\mathbf{u}_n)\Delta\mathbf{W}_n, \tag{1.24}$$

where $\phi(\Delta t A) = A^{-1}(e^{\Delta t A} - I)$, with $I \in \mathbb{R}^{d\times d}$ being the identity matrix.

In Figure 1.4, we use (1.24) to solve a one-dimensional Itô SDE with multiplicative noise. The figure was generated using Algorithm 7.3.5 of the Appendix. SETD converges strongly with order 0.5, as it can be seen in Figure 1.5. Later on, we review an analogue of the above numerical scheme, derived for Stratonovich-interpreted SDEs.

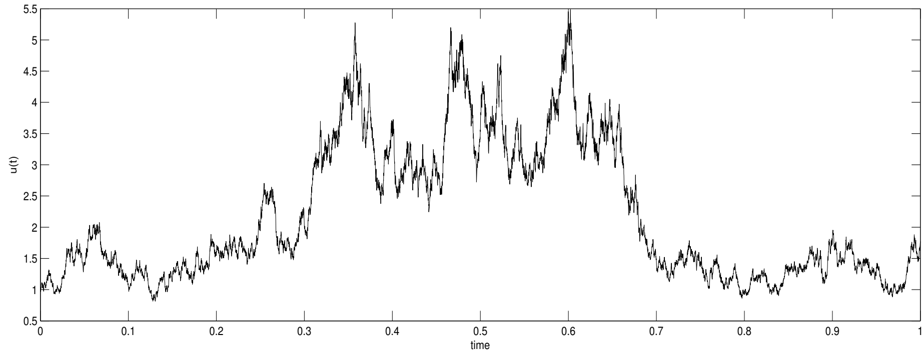


Figure 1.4: Approximate solution of $du = \big(Au + f(u)\big)dt + g(u)dW$ with $f = u$, $g(u) = -2u$, $A = 1$, after 10^4 iterations and with initial condition $u_0 = 1$.

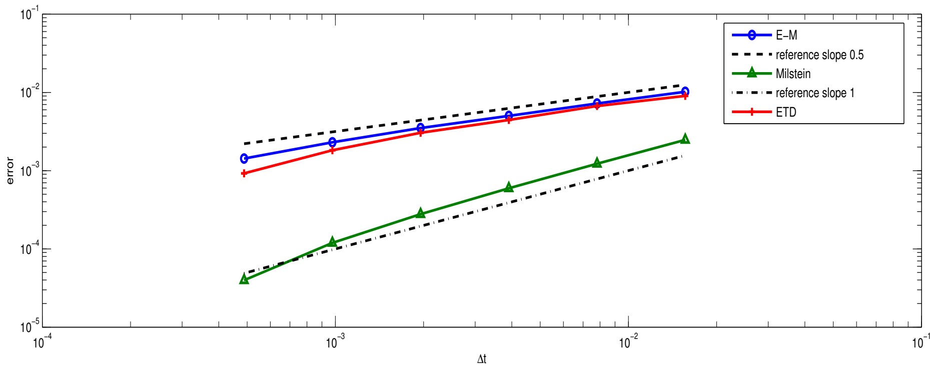


Figure 1.5: Orders of strong convergence for the Itô schemes Euler-Maruyama, Milstein and ETD. $M = 200$ samples, reference step-size 2^{-12}.

1.2.4 Exponential Milstein-type scheme for Itô SDEs

In this section we see an exponential analog of Milstein's scheme for Itô SDEs. We review the derivation of the scheme, following the approach of [37]. Based on this idea, we derive a new exponential-based Milstein-type scheme for Stratonovich SDEs later in Chapter 2.

Under Assumptions 1.1.9 and 1.1.10, we consider an Itô SODE of the same form as in equation (1.22) with solution given by

$$\mathbf{u}(t) = e^{(t-r)A}\mathbf{u}(r) + \int_r^t e^{(t-s)A}\mathbf{F}(\mathbf{u}(s))ds + \int_r^t e^{(t-s)A}G(\mathbf{u}(s))d\mathbf{W}(s).$$

In order to derive the higher order scheme, we firstly Talylor expand $\mathbf{F}(\mathbf{u}(s))$ and

$G(\mathbf{u}(s))$ around $\mathbf{u}(r)$, to obtain

$$\mathbf{u}(t) = e^{(t-r)A}\mathbf{u}(r) + \int_r^t e^{(t-s)A}\mathbf{F}(\mathbf{u}(r))ds + \int_r^t e^{(t-s)A}G(\mathbf{u}(r))d\mathbf{W}(s) \tag{1.25}$$

$$+ \int_r^t e^{(t-s)A}DG(\mathbf{u}(r))\big(\mathbf{u}(s) - \mathbf{u}(r)\big)d\mathbf{W}(s) + R_T,$$

where R_T stands for a higher order term that comes from the Taylor expansion. By the Euler-Maruyama scheme we have that

$$\mathbf{u}(s) = \mathbf{u}(r) + \mathbf{F}(\mathbf{u}(r))(r-s) + G(\mathbf{u}(r))\int_r^s d\mathbf{W}(\tau) + R_E,$$

where R_E denotes the remainder term of the Euler-Maruyama approximation. If we substitute for the difference $\mathbf{u}(s) - \mathbf{u}(r)$ in equation (1.25) above and we 'drop out' the remainder terms R_T, R_E, we get

$$\mathbf{u}(t) = e^{(t-r)A}\mathbf{u}(r) + \int_r^t e^{(t-s)A}\mathbf{F}(\mathbf{u}(r))ds + \int_r^t e^{(t-s)A}G(\mathbf{u}(r))d\mathbf{W}(s)$$

$$+ \int_r^t e^{(t-s)A}DG(\mathbf{u}(r))\Big(G(\mathbf{u}(r))\int_r^s d\mathbf{W}(\tau)\Big)d\mathbf{W}(s). \tag{1.26}$$

Note that apart from the remainder terms R_T, R_E we also 'drop out' the term $\mathbf{F}(\mathbf{u}(r))(r-s)$ which corresponds to the approximation of the drift term. The reason for this is that, as in the Itô Milstein scheme of Subsection 1.2.2, this term does not really contribute anything towards achieving a higher order. Then, the approximate solution as given by equation (1.26), considered in the time intervals $[t_n, t_{n+1}]$ reads

$$\mathbf{u}(t_{n+1}) = e^{\Delta tA}\mathbf{u}(t_n) + \mathbf{F}(\mathbf{u}(t_n))\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}ds + G(\mathbf{u}(t_n))\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}d\mathbf{W}(s)$$

$$+ \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}DG(\mathbf{u}(t_n))\Big(G(\mathbf{u}(t_n))\int_{t_n}^s d\mathbf{W}(\tau)\Big)d\mathbf{W}(s). \tag{1.27}$$

Finally, we write (1.27) in a component-wise form, so the k-th component of the higher order exponential Milstein-type scheme is given by

$$u_{k,n+1} = \sum_{q=1}^d a_{kq}u_q + \sum_{q=1}^d a_{kq}\mathbf{F}_q(\mathbf{u}_n) + \sum_{q=1}^d\sum_{j=1}^m a_{kq}G_{qj}(\mathbf{u}_n)\Delta W_{j,n}$$

$$+\frac{1}{2}\sum_{q=1}^d\sum_{i=1}^m\sum_{\ell=1}^d a_{kq}\frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell i}(\mathbf{u}_n)\big(\Delta W_{i,n}^2 - \Delta t\big)$$

$$+\frac{1}{2}\sum_{q=1}^{d}\sum_{i<j=1}^{m}\sum_{\ell=1}^{d}a_{kq}\Big(\frac{\partial G_{qj}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell i}(\mathbf{u}_n)+\frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell j}(\mathbf{u}_n)\Big)\Delta W_{i,n}\Delta W_{j,n}$$

$$+\frac{1}{2}\sum_{q=1}^{d}\sum_{i<j=1}^{m}\sum_{\ell=1}^{d}a_{kq}\Big(\frac{\partial G_{qj}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell i}(\mathbf{u}_n)+\frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell j}(\mathbf{u}_n)\Big)A_{ij,n}, \tag{1.28}$$

where a_{kq} are the kq-elements of the operator $e^{\Delta tA}$. The double integral of the fourth term of (1.27) has been expanded as in Example 4 of the Appendix, more details about the derivation can be found in [54]. Note that, the scheme as given in (1.28), approximates solutions for systems of Itô-interpreted SDEs with general noise i.e. non-diagonal, non-commutative, multiplicative. In case of diagonal noise, the non-diagonal terms cancel out, hence the scheme reduces to

$$u_{k,n+1}=\sum_{q=1}^{d}a_{kq}u_q+\sum_{q=1}^{d}a_{kq}\mathbf{F}_q(\mathbf{u}_n)+\sum_{q=1}^{d}a_{kq}G_{kk}(\mathbf{u}_n)\Delta W_{k,n}$$

$$+\frac{1}{2}\sum_{q=1}^{d}a_{kq}\frac{\partial G_{kk}}{\partial u_k}(\mathbf{u}_n)G_{kk}(\mathbf{u}_n)\big(\Delta W_{k,n}^2-\Delta t\big). \tag{1.29}$$

Moreover, in case of commutative noise, i.e. $\frac{\partial G_{qj}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell i}(\mathbf{u}_n)=\frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell j}(\mathbf{u}_n)$, the scheme is simplified to

$$u_{k,n+1}=\sum_{q=1}^{d}a_{kq}u_q+\sum_{q=1}^{d}a_{kq}\mathbf{F}_q(\mathbf{u}_n)+\sum_{q=1}^{d}\sum_{j=1}^{m}a_{kq}G_{qj}(\mathbf{u}_n)\Delta W_{j,n}$$

$$+\frac{1}{2}\sum_{q=1}^{d}\sum_{i=1}^{m}\sum_{\ell=1}^{d}a_{kq}\frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell i}(\mathbf{u}_n)\big(\Delta W_{i,n}^2-\Delta t\big)$$

$$+\frac{1}{2}\sum_{q=1}^{d}\sum_{i<j=1}^{m}\sum_{\ell=1}^{d}a_{kq}\Big(\frac{\partial G_{qj}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell i}(\mathbf{u}_n)+\frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell j}(\mathbf{u}_n)\Big)\Delta W_{i,n}\Delta W_{j,n}. \tag{1.30}$$

To sum up, in the cases of diagonal and commutative noise we end up with reduced versions of the original scheme, as given in (1.28) for general noise. The reduced versions (1.29) and (1.30), both benefit from not containing the double integral-terms, the simulation of which turns out to be quite challenging.

This Milstein-type scheme is known to converge faster than both the Euler-Maruyama scheme and the Itô exponential integrators scheme of Subsection 1.2.3. Its order of convergence is approximately $O(\Delta t)$, as it can be seen in Figure 1.6, which makes it more efficient than the previously mentioned schemes especially in case of commutative noise where the double integral computation can be avoided.

An extension of the scheme for SPDEs as well as the proof of strong convergence

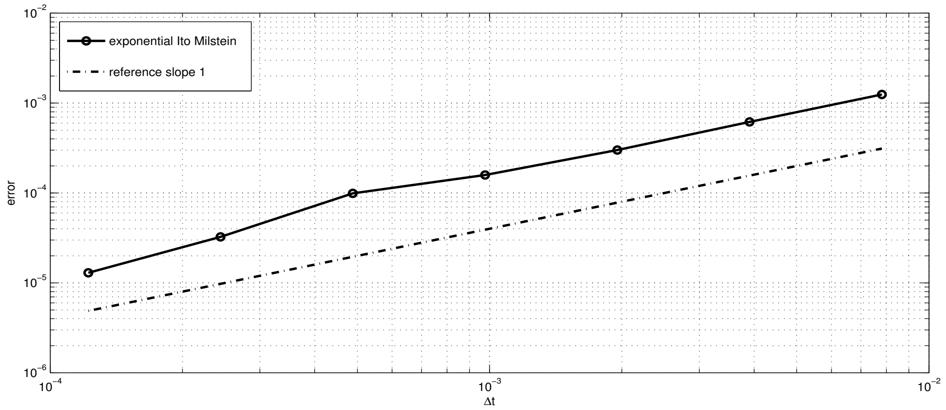


Figure 1.6: Itô exponential Milstein for 3-d Langevin system as given in (2.20). $M = 100$ samples, general i.e. non-diagonal, non-commutative noise. Noise intensity $\sigma = 0.1$, reference step-size 2^{-14}.

can be found in [37]. We examine the rate and prove strong convergence for a Stratonovich equivalent of this higher order SDEs scheme, in Section 2.3 of Chapter 2.

The numerical schemes that have been reviewed so far, concern approximating solutions to Itô SDEs. However, one could use the Itô-Stratonovich correction formula to convert a Stratonovich SDE into an equivalent Itô and then apply any of the above or other Itô numerical scheme for solving it. Alternatively, we can use numerical schemes that have been derived for and apply to Stratonovich-interpreted SDEs. The following numerical scheme, called Heun's method is one of them.

1.2.5 The stochastic Heun method

Heun's method is a specific case of a Runge-Kutta numerical scheme. The idea behind the deterministic scheme in order to approximate a point of a solution, is the following. Predict the left-end point tangent which underestimates the solution, then use Euler's method to predict the right-end point and hence predict the right-end point tangent. Next, find the line passing through the left-end point and having the same slope as the right-end point tangent; this overestimates the solution. Finally, take the line that starts at the left-end point and has the average slope of the tangent that underestimates the solution and the line that overestimates it.

We can now see the stochastic version of Heun's method and in order to do this let us consider the d-dimensional Stratonovich SDE of equation (1.10). Then, the stochastic Heun scheme, as derived in [42], reads

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \frac{1}{2}\big(\mathbf{F}(\mathbf{u}_n) + \mathbf{F}(\tilde{\mathbf{u}}_{n+1})\big)\Delta t + \frac{1}{2}\big(G(\mathbf{u}_n) + G(\tilde{\mathbf{u}}_{n+1})\big)\Delta \mathbf{W}_n. \tag{1.31}$$

By $\tilde{\mathbf{u}}_{n+1}$, we simply denote the point $\mathbf{u}_{n+1}$, which in the literature is often called predictor or supporting value. The prediction step calculates a rough approximation of the solution, typically using an explicit method. The corrector step on the other hand, uses the initial guess to perform the iteration of the implicit method and complete the approximation. Using this approach, one aims at gaining an improvement of the accuracy over an explicit scheme. However, the price that has to be paid lies in the computational effort required for estimating $\tilde{\mathbf{u}}_{n+1}$ and evaluating $G(\tilde{\mathbf{u}}_{n+1})$. For Heun's method, $\tilde{\mathbf{u}}_{n+1}$ is taken from the Euler-Maruyama scheme in the following way

$$\tilde{\mathbf{u}}_{n+1} = \mathbf{u}_n + \mathbf{F}(\mathbf{u}_n)\Delta t + G(\mathbf{u}_n)\Delta \mathbf{W}_n. \tag{1.32}$$

We work later in Chapter 3 on a generalised version of another predictor-corrector type scheme for which $\tilde{\mathbf{u}}_{n+1}$ may be taken from any explicit numerical scheme and we show how the choice of $\tilde{\mathbf{u}}_{n+1}$ affects the overall performance.

It is proved in [75] that Heun's method converges with a strong order of $O(\Delta t)$. However, in preparation for the proofs of Chapters 2 and 3, we give the basic steps of our approach to the strong convergence proof in the simplified one dimensional case.

As explained in the strong convergence section, we are going to seek for an upper bound for the difference between the exact and the approximate solution. In order to do this we need a sort of preparatory step which involves writing both solutions in a form that is convenient for our analysis. We start from the exact solution of the one-dimensional Stratonovich SDE of the form (1.8) which considered in the time intervals $[t_n, t_{n+1}]$ satisfies the integral equation

$$u(t_{n+1}) = u(t_n) + \int_{t_n}^{t_{n+1}} f(u(s))ds + \int_{t_n}^{t_{n+1}} g(u(s)) \circ dW(s). \tag{1.33}$$

By the Itô-Stratonovich correction formula, (1.33) is equivalent to

$$u(t_{n+1}) = u(t_n) + \int_{t_n}^{t_{n+1}} f(u(s)) + \int_{t_n}^{t_{n+1}} g(u(s))dW(s) + \frac{1}{2}\int_{t_n}^{t_{n+1}} g(u(s))g'(u(s))ds. \tag{1.34}$$

Then, (1.34) is finally written as the following recurrence

$$u(t_n) = u_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} f(u(s))ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} g(u(s))dW(s)$$
$$+ \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} g(u(s))g'(u(s))ds. \tag{1.35}$$

As regards the approximate solution, we start from rewriting the equivalent one-dimensional case of (1.31) in the following way

$$u_{n+1} = u_n + \frac{1}{2}\int_{t_n}^{t_{n+1}} \big(f(u_n) + f(\tilde{u}_{n+1})\big)ds + \frac{1}{2}\int_{t_n}^{t_{n+1}} \big(g(u_n) + g(\tilde{u}_{n+1})\big)dW(s). \quad (1.36)$$

Then, we substitute for the predictor $\tilde{u}_{n+1}$ and we Taylor expand $f(\tilde{u}_{n+1})$ and $g(\tilde{u}_{n+1})$ about u_n to get

$$f(\tilde{u}_{n+1}) = f(u_n) + f(u_n)f'(u_n)\Delta t + R_f,$$

$$g(\tilde{u}_{n+1}) = g(u_n) + g(u_n)g'(u_n)\Delta W_n + R_g$$

where $R_f := g(u_n)f'(u_n)\Delta W_n + h.o.t$, $R_g := f(u_n)g'(u_n)\Delta t + h.o.t.$ So, (1.36) can be rewritten as the following recurrence

$$\begin{aligned} u_n = u_0 &+ \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} f(u_k)ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} g(u_k)dW(s) \quad (1.37)\\ &+ \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} f(u_k)f'(u_k)\Delta t ds + \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} g(u_k)g'(u_k)dW^2(s). \end{aligned}$$

Now, we consider the difference between the exact and the approximate solution i.e. (1.35)-(1.37) and we group it in the following terms.

$$I := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \big(f(u(s)) - f(u_k)\big)ds$$

$$II := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \big(g(u(s)) - g(u_k)\big)dW(s)$$

$$III := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \big(g(u(s))g'(u(s)) - g(u_k)g'(u_k)\big)ds$$

$$IV := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} f(u_k)f'(u_k)\Delta t ds + R_f + R_g.$$

Note that the initial conditions are not included in any of the above terms. This is because we assume that there is no noise in the initial data, hence the two terms cancel out. Moreover, for term III we have used the rule provided in Lemma 1.4.1.

Then, we look at the mean square error as given by

$$\|u(t_n) - u_n\|_{L^2(\Omega,\mathbb{R})} = \Big(\mathbb{E}\big[\|u(t_n) - u_n\|_2^2\big]\Big)^{1/2}.$$

By substituting for the difference $u(t_n) - u_n$ we have

$$\begin{aligned}\|u(t_n) - u_n\|_{L^2(\Omega,\mathbb{R})} &= \Big(\mathbb{E}\big[\|I + II + III + IV\|_2^2\big]\Big)^{1/2} \\ &\leqslant 2\Big(\mathbb{E}[\|I\|_2^2] + \mathbb{E}[\|II\|_2^2] + \mathbb{E}[\|III\|_2^2] + \mathbb{E}[\|IV\|_2^2]\Big)^{1/2}.\end{aligned}$$

Then, we can work out the individual estimates and at the end put them together in order to obtain the final estimate. Terms $\mathbb{E}[\|I\|_2^2]$ and $\mathbb{E}[\|III\|_2^2]$ may be handled in a similar way. That is, use the Lipschitz condition for f and gg' as given in Assumptions 1.1.9 and 2.1.2 to show that the two terms are bounded by $\displaystyle\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(s) - u_k\|_2^2\big]ds$. For the term $\mathbb{E}[\|II\|_2^2]$, we additionally use Itô's Isometry, as seen in Proposition 1.1.6, to derive the same upper bound. Finally, the term $\mathbb{E}[\|IV\|_2^2]$ turns out to be bounded by Δt^2. By putting everything together, using Gronwall's Lemma and taking the square root, we obtain strong order of $O(\Delta t)$.

In Figures 1.7, 1.8 we see a one-dimensional and a two-dimensional example, both simulated by using Heun's scheme. Heun's function-file for Matlab can be found in Algorithm 7.3.4 in Appendix.

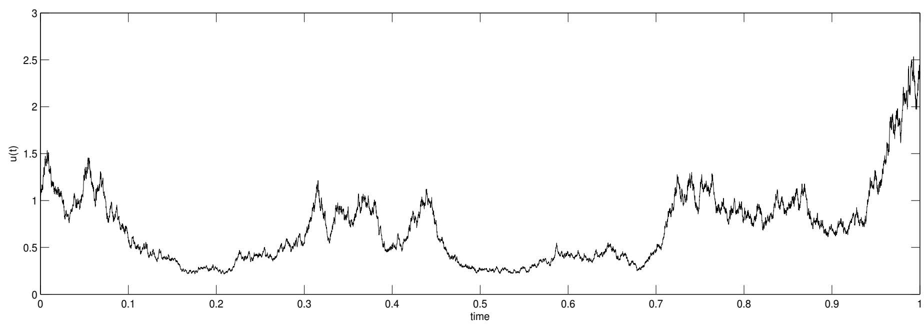


Figure 1.7: Solution approximated by 10^4 iterations of Heun's scheme. $f(u) = u$, $g(u) = 2u$, initial condition $u_0 = 1$.

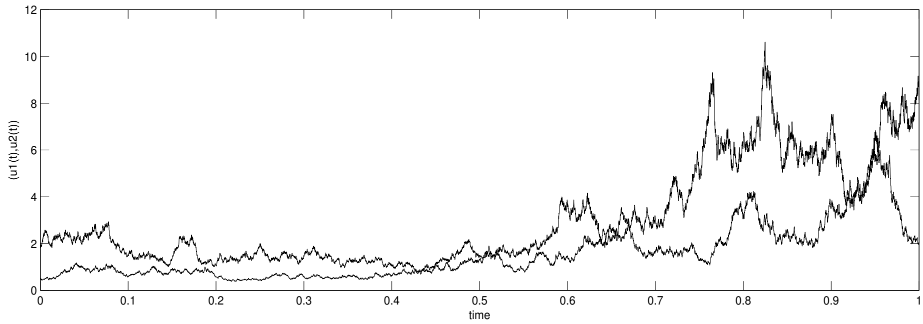


Figure 1.8: Approximate solution of a two-dimensional Stratonovich SDE with $\mathbf{F}(\mathbf{u}) = \mathbf{u}$ and $G(\mathbf{u}) = \left(\begin{smallmatrix} 2u_1 & 0 \\ 0 & 2u_2 \end{smallmatrix}\right)$. The initial condition is $(u_1(0), u_2(0)) = (0.5, 2)$.

1.2.6 Stratonovich exponential integrators

The following, is another numerical scheme for approximating solutions of Stratonovich SDEs. The scheme, firstly introduced in [1], is called the Stratonovich Exponential Integrators (SEI) and is a method that is based on the definition of the Stratonovich integral and also that, in a similar way to the Exponential Integrators for Itô, involves exponential functions derived by using the variation of constants formula.

In this section, we just review the derivation of SEI. In Chapter 2 however, we illustrate some numerical examples of the performance of the scheme as regards the errors and the order of convergence. Additionally, we investigate analytically its strong convergence in the mean square sense, that is by considering the expectation of the difference between the exact and the approximate solution squared.

So, let us consider a semilinear Stratonovich SDE of the form (1.10), with drift given by $A\mathbf{u} + \mathbf{F}(\mathbf{u})$

$$d\mathbf{u} = \big(A\mathbf{u} + \mathbf{F}(\mathbf{u})\big)dt + G(\mathbf{u}) \circ d\mathbf{W}, \tag{1.38}$$

with some initial condition $\mathbf{u}(0) = \mathbf{u}_0$, with A being a bounded linear operator and with $\mathbf{F}$, G satisfying Assumptions 1.1.9 and 1.1.10 as before. Then, under these Assumptions and after applying the variation of constants formula, (1.38) has a unique solution which satisfies the following integral equation

$$\mathbf{u}(t) = e^{tA}\mathbf{u}_0 + \int_0^t e^{(t-s)A}\mathbf{F}(\mathbf{u}(s))ds + \int_0^t e^{(t-s)A}G(\mathbf{u}(s)) \circ d\mathbf{W}(s), \tag{1.39}$$

or the following equation considered in the time intervals $[t_n, t_{n+1}]$

$$\mathbf{u}(t_{n+1}) = e^{\Delta tA}\mathbf{u}(t_n) + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{F}(\mathbf{u}(s))ds + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}G(\mathbf{u}(s)) \circ d\mathbf{W}(s) \tag{1.40}$$

Similarly with SETD, in order to derive a scheme that approximates the solution of (1.38), we evaluate $\mathbf{F}(\mathbf{u}(s))$ at the point $\mathbf{u}(t_n)$, we compute the deterministic integral and we approximate the stochastic integral by evaluating the function G at the mid-point of the integration interval. In particular, the deterministic integral is given by

$$\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{F}(\mathbf{u}(t_n))ds = A^{-1}(e^{\Delta tA} - I)\mathbf{F}(\mathbf{u}(t_n)) \tag{1.41}$$

and the stochastic integral, as suggested by the definition of Stratonovich integral, is approximated by

$$\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}G(\mathbf{u}(s)) \circ d\mathbf{W}(s) \approx \frac{1}{2}e^{\frac{\Delta tA}{2}}\big(G(\mathbf{u}(t_n)) + G(\mathbf{u}(t_{n+1}))\big)\Delta\mathbf{W}_n, \tag{1.42}$$

where $\Delta\mathbf{W}_n$ denotes the Brownian increments $\mathbf{W}(t_{n+1}) - \mathbf{W}(t_n)$. Then, we substitute (1.41) and (1.42) in (1.40) to get

$$\mathbf{u}_{n+1} = e^{\Delta tA}\mathbf{u}_n + \phi(\Delta tA)\mathbf{F}(\mathbf{u}_n) + \frac{1}{2}e^{\frac{\Delta tA}{2}}\big(G(\mathbf{u}_n) + G(\tilde{\mathbf{u}}_{n+1})\big)\Delta\mathbf{W}_n, \tag{1.43}$$

where $\phi(\Delta tA) = A^{-1}(e^{\Delta tA} - I)$ and $\tilde{\mathbf{u}}_{n+1}$ is a predictor given for example by Euler-Maruyama as in (1.32). Note that (1.43) is the form of the numerical scheme that can be implemented for numerical use, later in Chapter 2, we also use a different form of (1.43) which is more suitable for the analysis that we do.

The numerical experiments that have been carried out in [1] suggest that the SEI scheme converges with a strong order of approximately $O(\sqrt{\Delta t})$. We confirm this result in Chapter 2 where we give a proof of strong convergence.

1.3 Numerical comparisons

After reviewing some numerical schemes for approximating solutions of either Itô or Stratonovich SDEs, in this section, we compare the results that we get when using the previously presented schemes. Note that in the cases where we compare a scheme originally derived for Itô with a scheme for Stratonovich, we have to consider a corrected version of one of the two in order to reassure that we are comparing solutions of the same problem. In Figures 1.9 - 1.12, we plot on the left hand side the solution of the SDE approximated by two different schemes each time. On the right hand side, we plot the absolute difference between the approximations given by the two schemes on a logarithmic scale. This, indicates that the approximations provided by the previously mentioned numerical schemes agree at least to some extent. For all the comparisons, we have considered the same multiplicative noise, one-dimensional toy-example which we approximate in the interval $[0, 1]$ by performing 10^4 iterations. The initial conditions are $u_0 = 1$ and the functions used are $f(u) = u$, $g(u) = -2u$

for all the cases.

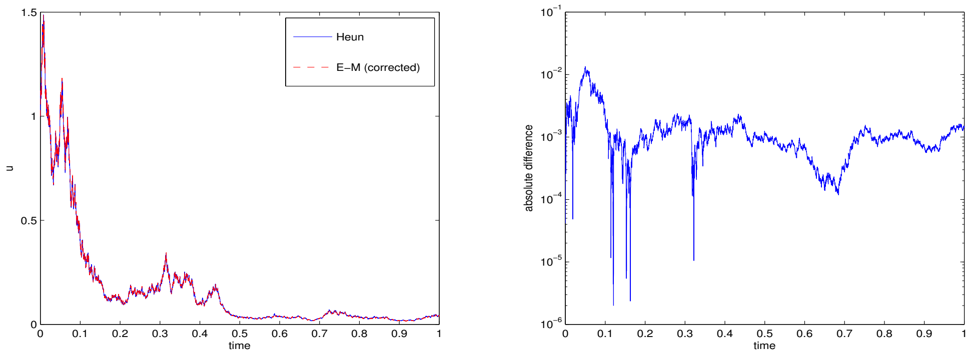


Figure 1.9: Comparison between corrected Euler-Maruyama and stochastic Heun schemes for approximating a Stratonovich SDE with $f(u) = u$, $g(u) = -2u$, $u_0 = 1$.

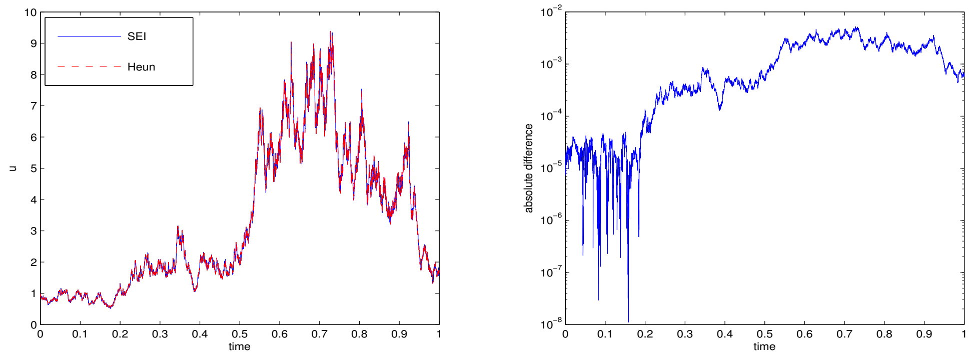


Figure 1.10: Comparison between Heun and Stratonovich Exponential Integrators for approximating a Stratonovich SDE with $f(u) = u$, $g(u) = -2u$ and $u_0 = 1$.

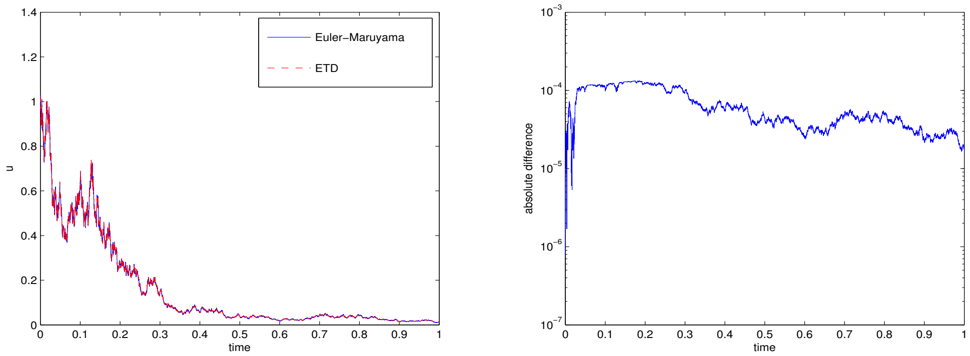


Figure 1.11: Comparison between Euler-Maruyama and ETD for approximating the equivalent Itô SDE of Figures 1.9 and 1.10.

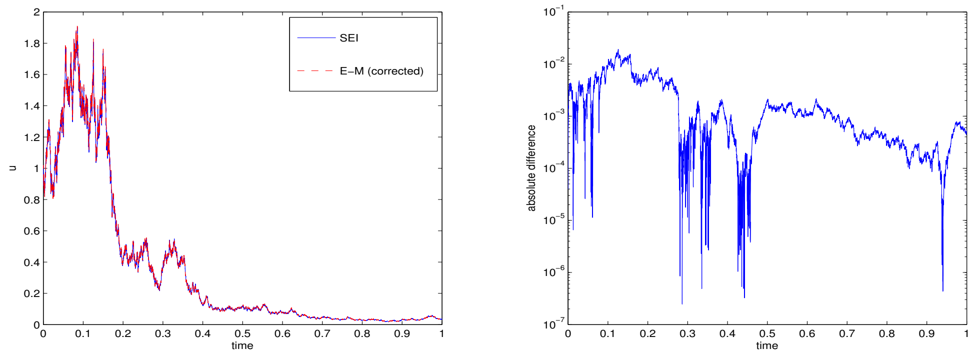


Figure 1.12: Corrected Euler-Maruyama compared to Exponential Integrators for Stratonovich for approximating an SDE with $f(u) = u$, $g(u) = -2u$ and $u_0 = 1$.

1.4 Some useful tools

In the last section of this chapter, we give a collection of Lemmas and other tools that turn out to be useful for our analysis throughout this thesis.

We start from the following result which just states what is the variance of the Brownian increment of zero mean. A sketch of the proof for the following lemma can be found in Section 7.2 of the Appendix.

Lemma 1.4.1. *[15]*

Let $\{W(t) : t \geqslant 0\}$ be a Wiener process and consider an SDE of the form (1.1). Then

$$\mathbb{E}[(dW_t)^2] = dt.$$

One of the inequalities that we use later on in order to obtain estimates for several terms is Gronwall's Inequality. Given that a function satisfies a specific integral inequality, Gronawall's Inequality or Lemma provides an upper bound for the function.

Lemma 1.4.2. ***Gronwall's Inequality*** *[42, 58]*
Let $\alpha, \beta : [t_0, T] \to \mathbb{R}$ integrable with

$$0 \leqslant \alpha(t) \leqslant \beta(t) + L \int_{t_0}^{t} \alpha(s) ds, \;\; t \in [t_0, T],\, L > 0$$

then

$$\alpha(t) \leqslant \beta(t) + L \int_{t_0}^{t} e^{L(t-s)} \beta(s) ds, \;\; t \in [t_0, T].$$

Lemma 1.4.3. ***Jensen's Inequality*** *[54, 63]*
Let $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function and X be a real valued random variable with

$\mathbb{E}(X) < \infty$. Then

$$\phi\big(\mathbb{E}(X)\big) \leqslant \mathbb{E}\big(\phi(X)\big).$$

In particular, for $p > 1$:

$$\big(\mathbb{E}|X|\big)^p \leqslant \mathbb{E}\big(|X|^p\big).$$

Theorem 1.4.4. ***Fubini's Theorem*** *[54, 74]*

Let $(X, \mathcal{F}_1, \mu)$ and $(Y, \mathcal{F}_2, \lambda)$ be σ-finite measure spaces and let $u : X \times Y \to Z$ be a measurable function such that

$$\int_Y \Big(\int_X \|u(x,y)\|_Z d\mu(x) \Big) d\lambda(y) < \infty.$$

Then, u is integrable with respect to the product measure $(\mu \times \lambda)$ and

$$\int_{X\times Y} u(x,y) d(\mu\times\lambda)(x,y) = \int_Y \Big(\int_X u(x,y) d\mu(x) \Big) d\lambda(y) = \int_X \Big(\int_Y u(x,y) d\lambda(y) \Big) d\mu(x).$$

In other words, under the conditions specified above, the double integral may be given by an expression of iterated integrals and the order of integration may be reversed.

Lemma 1.4.5. ***Hölder's Inequality*** *[9]*

Let $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$. Assume that u, v are measurable functions with respect to some measure μ and that $u \in L^p$ and $v \in L^q$. Then, $uv \in L^1$ and

$$\|uv\|_1 = \int |uv| d\mu \leqslant \|u\|_p \|v\|_q.$$

Lemma 1.4.6. ***Minkowski's Inequality*** *[85]*

Let $u, v \in L^p$, then

$$\|u + v\|_p \leqslant \|u\|_p + \|v\|_p.$$

So far we have considered Itô and Stratonovich semilinear SDEs of the form (1.22) and (1.38) respectively, where A was a bounded linear operator (matrix) in $\mathbb{R}^{d\times d}$. At this stage, in preparation for looking at SPDEs, we extend our definitions and consider A to be a linear operator on a Hilbert space H with norm $\|\cdot\|_H$. In what follows, we consider the semigroup approach which will later enable us to see an SPDE as an SDE on a suitable function space. We start from defining a semigroup.

Definition 1.4.7. ***Semigroup*** *[72]*

On a Hilbert space H we define a family of operators $\{S(t)\}_{t\geqslant 0}$ which satisfy the properties

$S(0) = I$

$S(t)S(s) = S(s)S(t) = S(s+t)\ \forall s, t \geqslant 0$ and

$S(t)$ is continuous $\forall t \geqslant 0$.

We call the family $\{S(t)\}_{t\geqslant 0}$ a C^0 semigroup.

Assumption 1.4.8. *We assume that A is a linear operator and that $(-A)$ generates an analytic semigroup e^{-tA}, $t > 0$.*

We further assume that A generates the semigroup of operators e^{tA}. Then, some properties of e^{tA} are given in the following Proposition

Proposition 1.4.9. ***Smoothing Properties of the semi-group*** *[20, 56, 66, 68]*

Let H be a Hilbert space and $L(H)$ the set of bounded linear mappings from H to H. Also let $\alpha > 0$, $\beta \geqslant 0$, $0 \leqslant \gamma \leqslant 1$. Then there exists $C > 0$ such that

$$\left\|(-A)^{\beta} S(t)\right\|_{L(H)} \leqslant C t^{-\beta}, \quad t > 0,$$

$$\left\|(-A)^{-\gamma}\big(I - S(t)\big)\right\|_{L(H)} \leqslant C t^{\gamma}, \quad t \geqslant 0.$$

In addition,

$$(-A)^{\beta} S(t) = S(t)(-A)^{\beta} \ \text{ on } \mathcal{D}((-A)^{\beta})$$

This section concludes the introductory chapter which provides the necessary background for studying numerical schemes for SDEs and establishes the framework within which we work. We start Chapter 2, by proving that the SEI scheme, reviewed in subsection 1.2.6, converges with a strong order of approximately $O(\sqrt{\Delta t})$. Later in Chapter 2, we suggest a new Milstein-type exponential scheme for Stratonovich SDEs and we prove that it converges strongly with a higher order than SEI.

Chapter 2

Stratonovich Exponential Integrators

2.1 Strong convergence of SEI in d dimensions

In this section we prove analytically strong convergence for the SEI scheme as given in equation (1.43). Let us consider the semilinear Stratonovich SDE of the form (1.38) in the bounded time interval $[0, T]$ with $t_0 = 0$ and final computational time T. We consider the case $d = m$ with multiplicative noise and with the same assumptions as in Chapter 1. Note that for additive noise, $G(\mathbf{u})$ does not depend on $\mathbf{u}$ and the Stratonovich SDE coincides with the Itô. This is very easy to see, simply recall the Itô-Stratonovich correction formula, as seen in equation (1.9), and check that the correction term becomes zero when $G(\mathbf{u})$ is constant.

Further to the assumptions that we impose on $\mathbf{F}$ and G in Chapter 1, we require that they also satisfy the following.

Assumption 2.1.1. *We assume that $\boldsymbol{F}(\boldsymbol{u})$ and $G(\boldsymbol{u})$ are twice continuously differentiable with uniformly bounded second derivatives.*

Assumption 2.1.2. *Let $\boldsymbol{c}(\boldsymbol{u}) = \big(c_1(\boldsymbol{u}), \ldots, c_d(\boldsymbol{u})\big) \in \mathbb{R}^d$ with $c_i(\boldsymbol{u}) = \displaystyle\sum_{j=1}^{m} \frac{\partial G_{ij}(\boldsymbol{u})}{\partial u_i}$. We assume that there exists a constant $C_l > 0$ such that $\boldsymbol{c}(\boldsymbol{u})$ satisfies*

$$\big\|\boldsymbol{c}(\boldsymbol{u}) - \boldsymbol{c}(\boldsymbol{v})\big\|_2^2 \leqslant C_l \big\|\boldsymbol{u} - \boldsymbol{v}\big\|_2^2, \quad \textit{for } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d.$$

Note that in this context, the Lipschitz condition implies linear growth of the function. The opposite however, is not true. Later on, when we consider SPDEs, the linear growth condition will be a useful tool in our analysis. Then, under 1.1.9, 1.1.10 and the extra Assumption 2.1.2, after applying the variation of constants formula, the mild solution of (1.38) satisfies the integral equation (1.39), or the following Itô equation considered in the time interval $[t_n, t_{n+1}]$, which is equivalent to (1.40) .

$$\mathbf{u}(t_{n+1}) = e^{\Delta t A}\mathbf{u}(t_n) + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{F}(\mathbf{u}(s))ds + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}G(\mathbf{u}(s))d\mathbf{W}(s)$$
$$+\frac{1}{2}\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{c}(\mathbf{u}(s))ds.$$

If we consider the above equation 'piecewisely' in the subintervals $[t_k, t_{k+1}), k = 0, \dots, n-1$, the exact solution is given by

$$\mathbf{u}(t_n) = e^{t_n A}\mathbf{u}_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{F}(\mathbf{u}(s))ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}G(\mathbf{u}(s))d\mathbf{W}(s)$$
$$+\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s))ds. \tag{2.1}$$

Note that, (2.1) is the so called 'corrected' version of the Stratonovich solution. The reason for which we consider this version is because it enables us to use the Itô isometry, an important tool for the analysis that follows.

Next, we derive a slightly different form of the approximate solution given by the SEI scheme in (1.43), to be used for the convergence proof. Recall that we distinguish between the exact and the approximate solution by using the notation $\mathbf{u}(t_n)$ and $\mathbf{u}_n$ respectively, when both are evaluated for example at the same time t_n. Looking back at (1.43), we start by expanding $G(\tilde{\mathbf{u}}_{n+1})$ in the following way. Let $\tilde{\mathbf{u}}_{n+1}$ be given by

$$\tilde{\mathbf{u}}_{n+1} = \mathbf{u}_n + \big(A\mathbf{u}_n + \mathbf{F}(\mathbf{u}_n)\big)\Delta t + G(\mathbf{u}_n)\Delta\mathbf{W}_n.$$

Then,

$$G(\tilde{\mathbf{u}}_{n+1}) = G\Big(\mathbf{u}_n + \big(A\mathbf{u}_n + \mathbf{F}(\mathbf{u}_n)\big)\Delta t + G(\mathbf{u}_n)\Delta\mathbf{W}_n\Big) := G(\mathbf{u}_n + \mathbf{h}),$$

where

$$\mathbf{h} := \big(A\mathbf{u}_n + \mathbf{F}(\mathbf{u}_n)\big)\Delta t + G(\mathbf{u}_n)\Delta\mathbf{W}_n. \tag{2.2}$$

Next, we Taylor expand $G(\mathbf{u}_n + \mathbf{h})$ about $\mathbf{u}_n$ in the interval $[t_n, t_{n+1}]$ to get

$$G(\mathbf{u}_n + \mathbf{h}) = G(\mathbf{u}_n) + DG(\mathbf{u}_n)[\mathbf{h}] + \mathbf{R}_1,$$

where $DG(\mathbf{u}) : \mathbb{R}^{d\times d} \to \mathbb{R}^{d^3}$ is a rank-three tensor as defined in Subsection 1.2.2, for more details see Section 7.1 of the Appendix. The second order remainder term $\mathbf{R}_1$, is given by the following formula in integral form

$$\mathbf{R}_1 = \frac{1}{2!}\int_0^1 (1-\theta)\mathbf{h}^\top D^2 G\big(\mathbf{u}_n + \theta\mathbf{h}\big)\mathbf{h}d\theta,$$

where $D^2G(\mathbf{u}) : \mathbb{R}^{d\times d} \to \mathbb{R}^{d^4}$ is a rank-four tensor, including up to second order partial derivatives, given by

$$D^2G(\mathbf{u}) = \begin{pmatrix} \nabla^2 G_{11}(\mathbf{u}) & \dots & \nabla^2 G_{1d}(\mathbf{u}) \\ \vdots & \ddots & \vdots \\ \nabla^2 G_{d1}(\mathbf{u}) & \dots & \nabla^2 G_{dd}(\mathbf{u}) \end{pmatrix},$$

with ij-elements

$$\nabla^2 G_{ij}(\mathbf{u}) = \nabla\big(\nabla G_{ij}(\mathbf{u})\big) = \nabla\Big(\frac{\partial G_{ij}(\mathbf{u})}{\partial u_1}, \dots, \frac{\partial G_{ij}(\mathbf{u})}{\partial u_d}\Big)^\top$$

and

$$\nabla\frac{\partial G_{ij}(\mathbf{u})}{\partial u_k} = \Big(\frac{\partial^2 G_{ij}(\mathbf{u})}{\partial u_k \partial u_1}, \dots, \frac{\partial^2 G_{ij}(\mathbf{u})}{\partial u_k^2}, \dots, \frac{\partial^2 G_{ij}(\mathbf{u})}{\partial u_k \partial u_2}\Big)^\top,$$

for fixed $i, j, k = 1, \dots, d$.

We can now substitute for $G(\tilde{\mathbf{u}}_{n+1})$, hence the third term of (1.43) is given by

$$\frac{1}{2}e^{\frac{\Delta tA}{2}}\big(G(\mathbf{u}_n) + G(\tilde{\mathbf{u}}_{n+1})\big)\Delta\mathbf{W}_n = \int_{t_n}^{t_{n+1}} e^{\frac{\Delta tA}{2}} G(\mathbf{u}_n)d\mathbf{W}(s)$$

$$+\frac{1}{2}\int_{t_n}^{t_{n+1}} e^{\frac{\Delta tA}{2}} DG(\mathbf{u}_n)[\mathbf{h}]d\mathbf{W}(s) + \frac{1}{2}\int_{t_n}^{t_{n+1}} e^{\frac{\Delta tA}{2}}\mathbf{R}_1 d\mathbf{W}(s).$$

Thus, the approximate solution can be rewritten as

$$\mathbf{u}_{n+1} = e^{\Delta tA}\mathbf{u}_n + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{F}(\mathbf{u}_n)ds + \int_{t_n}^{t_{n+1}} e^{\frac{\Delta tA}{2}} G(\mathbf{u}_n)d\mathbf{W}(s) \tag{2.3}$$

$$+\frac{1}{2}\int_{t_n}^{t_{n+1}} e^{\frac{\Delta tA}{2}}\mathbf{c}(\mathbf{u}_n)ds + \frac{1}{2}\int_{t_n}^{t_{n+1}} e^{\frac{\Delta tA}{2}} DG(\mathbf{u}_n)\mathbf{F}(\mathbf{u}_n)\Delta t d\mathbf{W}(s)$$

$$+\frac{1}{4}\int_{t_n}^{t_{n+1}} e^{\frac{\Delta tA}{2}}\Big(\int_0^1 (1-\theta)D^2G\big(\mathbf{u}(t_n) + \theta\mathbf{h}\big)[\mathbf{h}]^2 d\theta\Big)d\mathbf{W}(s),$$

where $\mathbf{c}(\mathbf{u}_n) \in \mathbb{R}^d$ is as defined in Assumption 2.1.2. This leads us to the following recurrence, which is the form of the approximate solution that we use for analysis.

Note that it includes the Taylor error from the prediction term.

$$\mathbf{u}_n = e^{t_n A}\mathbf{u}_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{F}(\mathbf{u}_k)ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}G(\mathbf{u}_k)d\mathbf{W}(s) \quad (2.4)$$

$$+\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}\hat{\mathbf{c}}(\mathbf{u}_k)ds + \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k)\Delta t d\mathbf{W}(s)$$

$$+\frac{1}{4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}\Big(\int_0^1 (1-\theta)\mathbf{h}^\top D^2G\big(\mathbf{u}_k+\theta\mathbf{h}\big)\mathbf{h}d\theta\Big)d\mathbf{W}(s).$$

The Stratonovich interpretation of the stochastic integral is reflected not only by the evaluation of the diffusion G at the midpoints, but also by the evaluation of the exponential term at the midpoint of each subinterval, as seen in the terms $e^{(t_n-t_{k+\frac{1}{2}})A}$.

To prove strong convergence for the Stratonovich SEI scheme, we consider the difference between (2.1) and (2.4) that is the difference between the exact and the approximate solution in the mean square sense. We consider

$$\big\|\mathbf{u}(t_n)-\mathbf{u}_n\big\|_{L^2(\Omega,\mathbb{R}^d)} = \Big(\mathbb{E}\big\|\mathbf{u}(t_n)-\mathbf{u}_n\big\|_2^2\Big)^{1/2}.$$

Then, the difference at the final approximation step is given by

$$\mathbf{u}(t_n)-\mathbf{u}_n = e^{t_n A}\mathbf{u}_0 - e^{t_n A}\mathbf{u}_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{F}(\mathbf{u}(s))ds - \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{F}(\mathbf{u}_k)ds \quad (2.5)$$

$$+\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}G(\mathbf{u}(s))d\mathbf{W}(s) - \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}G(\mathbf{u}_k)d\mathbf{W}(s)$$

$$+\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s))ds - \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}\hat{\mathbf{c}}(\mathbf{u}_k)ds$$

$$-\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k)\Delta t d\mathbf{W}(s)$$

$$-\frac{1}{4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}\left(\int_0^1 (1-\theta)\mathbf{h}^\top D^2G(\mathbf{u}_k+\theta\mathbf{h})\mathbf{h}\,d\theta\right)d\mathbf{W}(s).$$

As we did for the sketch of the proof for the Heun scheme in Subsection 1.2.5, we assume that there is no error in the initial data $\mathbf{u}_0$, thus when we consider the difference between the exact and the approximate solution, the initial condition terms cancel out. Next, (2.5) is grouped in the following terms

$$I := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\Big(\mathbf{F}(\mathbf{u}(s)) - \mathbf{F}(\mathbf{u}_k)\Big)ds$$

$$II := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \Big(e^{(t_n-s)A}G(\mathbf{u}(s)) - e^{(t_n-t_{k+\frac{1}{2}})A}G(\mathbf{u}_k)\Big)d\mathbf{W}(s)$$

$$III := \frac{1}{2}\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \Big(e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s)) - e^{(t_n-t_{k+\frac{1}{2}})A}\mathbf{c}(\mathbf{u}_k)\Big)ds$$

$$IV := -\frac{1}{2}\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k)\Delta t d\mathbf{W}(s)$$

$$-\frac{1}{4}\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}\left(\int_0^1 (1-\theta)\mathbf{h}^\top D^2G(\mathbf{u}_k+\theta\mathbf{h})\mathbf{h}\,d\theta\right)d\mathbf{W}(s).$$

Following the above notation, by the triangle inequality we have

$$\big\|\mathbf{u}(t_n) - \mathbf{u}_n\big\|^2_{L^2(\Omega,\mathbb{R}^d)} = \Big(\mathbb{E}\Big[\big\|I + II + III + IV\big\|_2^2\Big]\Big)^{1/2}$$

$$\leqslant 2\Big(\mathbb{E}\big[\|I\|_2^2\big] + \mathbb{E}\big[\|II\|_2^2\big] + \mathbb{E}\big[\|III\|_2^2\big] + \mathbb{E}\big[\|IV\|_2^2\big]\Big)^{1/2}.$$

Now, we consider each term separately and at the end we deduce the order of convergence by putting all the estimates together.

Lemma 2.1.3. *Let Assumptions 1.1.9 and 1.1.10 hold. Then*

$$\mathbb{E}\big[\|I\|_2^2\big] \leqslant C_1 \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|\boldsymbol{u}(s) - \boldsymbol{u}_k\|_2^2\big]ds,$$

where $C_1 = C_l^2 T \sup\limits_{t_0\leqslant s\leqslant T} \mathbb{E}\Big[\big\|e^{(T-s)A}\big\|_F^2\Big]$ *is a positive constant independent of* Δt *and* C_l *is the Lipschitz constant.*

Proof.
First of all

$$\mathbb{E}\big[\|I\|_2^2\big] = \mathbb{E}\Big[\Big\|\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\big(\mathbf{F}(\mathbf{u}(s)) - \mathbf{F}(\mathbf{u}_k)\big)ds\Big\|_2^2\Big].$$

By Hölder's Inequality 1.4.5, we get that

$$\mathbb{E}\big[\|I\|_2^2\big] \leqslant \mathbb{E}\Big[T\int_{t_0}^{T} \big\|e^{(T-s)A}\big(\mathbf{F}(\mathbf{u}(s)) - \mathbf{F}(\mathbf{u}_k)\big)\big\|_2^2 ds\Big].$$

Then, by considering the expectation as an integral, Fubini's Theorem 1.4.4 yields

$$\begin{aligned}\mathbb{E}\big[\|I\|_2^2\big] &\leqslant T\int_{t_0}^{T}\mathbb{E}\Big[\big\|e^{(T-s)A}\big(\mathbf{F}(\mathbf{u}(s))-\mathbf{F}(\mathbf{u}_k)\big)\big\|_2^2\Big]ds\\ &= T\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\Big[\big\|e^{(t_n-s)A}\big(\mathbf{F}(\mathbf{u}(s))-\mathbf{F}(\mathbf{u}_k)\big)\big\|_2^2\Big]ds.\end{aligned}$$

Using again Hölder's Inequality 1.4.5 we get that

$$\mathbb{E}\big[\|I\|_2^2\big] \leqslant T\sum_{k=0}^{n-1}\sup_{t_k\leqslant s\leqslant t_{k+1}}\mathbb{E}\Big[\big\|e^{(t_n-s)A}\big\|_F^2\Big]\int_{t_k}^{t_{k+1}}\mathbb{E}\Big[\big\|\mathbf{F}(\mathbf{u}(s))-\mathbf{F}(\mathbf{u}_k)\big\|_2^2\Big]ds,$$

which by Lipschitz Assumption 1.1.9 becomes

$$\leqslant C_l^2 T\sum_{k=0}^{n-1}\sup_{t_k\leqslant s\leqslant t_{k+1}}\mathbb{E}\Big[\big\|e^{(t_n-s)A}\big\|_F^2\Big]\int_{t_k}^{t_{k+1}}\mathbb{E}\Big[\big\|\mathbf{u}(s)-\mathbf{u}_k\big\|_2^2\Big]ds.$$

The fact that A is a bounded operator implies that the term $e^{(t_n-s)A}$ is also bounded as s varies between $[t_k, t_{k+1}]$. Hence, we have the result. □

The term $\mathbb{E}\big[\|II\|_2^2\big]$ is handled as two different terms, the first of which is bounded in a similar way with $\mathbb{E}\big[\|I\|_2^2\big]$. For the second term we need some further tools. We introduce the notation $e^{sA} := S(s)$ and based on the following Remark 2.1.4 and Proposition 1.4.9, we are going to find an upper bound for the second term of $\mathbb{E}\big[\|II\|_2^2\big]$.

Remark 2.1.4.

$$\begin{aligned}S(t_n-s)-S(t_n-t_{k+\frac{1}{2}}) &= S(t_n-s)(-A)^{\gamma}(-A)^{-\gamma}\big(I-S(s-t_{k+\frac{1}{2}})\big)\\ &= (-A)^{\gamma}S(t_n-s)(-A)^{-\gamma}\big(I-S(s-t_{k+\frac{1}{2}})\big),\end{aligned}$$

where $0\leqslant\gamma\leqslant 1$.

Lemma 2.1.5.
Let Assumptions 1.1.9 1.1.10 hold. Then

$$\mathbb{E}\big[\|II\|_2^2\big] \leqslant C_{2,1}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\big[\|\boldsymbol{u}(s)-\boldsymbol{u}_k\|_2^2\big]ds + C_{2,2}\Delta t\big(\ln(\Delta t)\big)^2,$$

where

$$C_{2,1} = 2C_l^2 T\sup_{0\leqslant s\leqslant T}\mathbb{E}\big[\|e^{(T-s)A}\|_F^2\big]$$

and

$$C_{2,2} = 2C^2 \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}\big[\|G(\boldsymbol{u}_k)\|_F^2\big]$$

are positive constants independent of Δt *and* $C > 0$ *constant as in Proposition 1.4.9.*

Proof.
We add and subtract $S(t_n - s)G(\mathbf{u}_k)$ to II and apply Itô's isometry to get

$$\mathbb{E}\Big[\|II\|_2^2\Big] = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|S(t_n-s)\big(G(\mathbf{u}(s))-G(\mathbf{u}_k)\big)+\big(S(t_n-s)-S(t_n-t_{k+\frac{1}{2}})\big)G(\mathbf{u}_k)\big\|_F^2\Big]ds$$

$$\leqslant 2\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|S(t_n - s)\big(G(\mathbf{u}(s)) - G(\mathbf{u}_k)\big)\big\|_F^2\Big]ds$$

$$+ 2\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|\big(S(t_n - s) - S(t_n - t_{k+\frac{1}{2}})\big)G(\mathbf{u}_k)\big\|_F^2\Big]ds.$$

So,

$$\mathbb{E}\Big[\|II\|_2^2\Big] \leqslant 2\big(II_1 + II_2\big), \tag{2.6}$$

where

$$II_1 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|S(t_n - s)\big(G(\mathbf{u}(s)) - G(\mathbf{u}_k)\big)\big\|_F^2\Big]ds$$

and

$$II_2 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|\big(S(t_n - s) - S(t_n - t_{k+\frac{1}{2}})\big)G(\mathbf{u}_k)\big\|_F^2\Big]ds.$$

An upper bound for the term II_1 is given by (2.7) below and is found in a similar way to the bound for $\mathbb{E}\big[\|I\|_2^2\big]$.

$$II_1 \leqslant C_{2,1} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|\mathbf{u}(s) - \mathbf{u}_k\|_2^2\big]ds, \tag{2.7}$$

where $C_{2,1} = 2C_l^2 T \sup\limits_{0 \leqslant s \leqslant T} \mathbb{E}\big[\|e^{(T-s)A}\|_F^2\big]$ and C_l is a Lipschitz constant.

We now obtain an upper bound on the term II_2. By Hölder's inequality, Remark 2.1.4 with $\gamma = 1$ and Proposition 1.4.9, we get that

$$II_2 \leqslant C^2 \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}\big[\|G(\mathbf{u}_k)\|_F^2\big] \Big(\Delta t^2 + \sum_{k=0}^{n-2} \int_{t_k}^{t_{k+1}} \Big(\frac{s - t_{k+\frac{1}{2}}}{t_n - s}\Big)^2 ds\Big)$$

with $C > 0$ constant as in Proposition 1.4.9.

Since s takes its values from the interval $[t_k, t_{k+1}]$, we have that $\frac{1}{t_n - s} \leqslant \frac{1}{t_n - t_{k+1}}$.

Hence,

$$II_2 \leqslant C^2 \sup_{0\leqslant k\leqslant n-1} \mathbb{E}\big[\|G(\mathbf{u}_k)\|_F^2\big]\Big(\Delta t^2 + \sum_{k=0}^{n-2} \frac{1}{(t_n - t_{k+1})^2} \int_{t_k}^{t_{k+1}} (s - t_{k+\frac{1}{2}})^2 ds\Big)$$

$$\leqslant C^2 \sup_{0\leqslant k\leqslant n-1} \mathbb{E}\big[\|G(\mathbf{u}_k)\|_F^2\big]\Big(\Delta t^2 + \Delta t \sum_{k=0}^{n-2} \frac{\Delta t^2}{(t_n - t_{k+1})^2}\Big)$$

$$\leqslant C_{2,2}\Delta t\big(\ln(\Delta t)\big)^2, \tag{2.8}$$

with $C_{2,2} = C^2 \sup_{0\leqslant k\leqslant n-1} \mathbb{E}\big[\|G(\mathbf{u}_k)\|_F^2\big]$. Finally, combining (2.6) with (2.7) and (2.8) finishes the proof.

□

For the term III we work in a similar way as for II. That is we apply Itô's isometry and Minkowski's inequality to get two terms which we bound separately in the following lemma.

Lemma 2.1.6.
Let Assumptions 1.1.9, 1.1.10 and 2.1.2 hold. Then

$$\mathbb{E}\big[\|III\|_2^2\big] \leqslant C_{3,1} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|\boldsymbol{u}(s) - \boldsymbol{u}_k\|_2^2\big] ds + C_{3,2}\Delta t\big(\ln(\Delta t)\big)^2,$$

where

$$C_{3,1} = C_l^2 \sup_{0\leqslant s\leqslant T} \mathbb{E}\big[\|e^{(T-s)A}\|_F^2\big],$$

$$C_{3,2} = C^2 \sup_{0\leqslant k\leqslant n-1} \mathbb{E}\big[\|G(\boldsymbol{u}_k)\|_F^2\big]$$

are positive constants independent of Δt *and* C_l, C *are the Lipschitz and the constant from Proposition 1.4.9 respectively.*

Proof.
Follow the same steps as the proof of Lemma 2.1.5 to get that

$$\mathbb{E}\Big[\big\|III\big\|_2^2\Big] \leqslant III_1 + III_2, \tag{2.9}$$

where

$$III_1 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|e^{(t_n-s)A}\big(\mathbf{c}(\mathbf{u}(s)) - \mathbf{c}(\mathbf{u}_k)\big)\big\|_2^2\Big] ds$$

and

$$III_2 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|\big(e^{(t_n-s)A} - e^{(t_n-t_{k+\frac{1}{2}})A}\big)\mathbf{c}(\mathbf{u}_k)\big\|_2^2\Big] ds.$$

Then the upper bounds for the terms III_1 and III_2, are given by the following equations.

$$III_1 \leqslant C_{3,1} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|\mathbf{u}(s) - \mathbf{u}_k\|_2^2\big] ds \tag{2.10}$$

and

$$III_2 \leqslant C_{3,2} \Delta t \big(\ln(\Delta t)\big)^2. \tag{2.11}$$

Combining (2.9) with (2.10) and (2.11) finishes the proof. □

Finally, the upper bound for the fourth term is given in the following lemma.

Lemma 2.1.7.
Under the Assumptions 1.1.9, 1.1.10 and the Assumption 2.1.2

$$\mathbb{E}\big[\|IV\|_2^2\big] \leqslant C_{4,1}\Delta t^3 + C_{4,2}\big(\Delta t^{7/2} + \Delta t^4\big),$$

where

$$C_{4,1} = \sup_{0\leqslant k\leqslant n-1} \mathbb{E}\big[\|DG(\boldsymbol{u}_k)\boldsymbol{F}(\boldsymbol{u}_k))\|_F^2\big] \sup_{0\leqslant k\leqslant n-1} \mathbb{E}\big[\|e^{(T-t_{k+\frac{1}{2}})A}\|_F^2\big] ds$$

and $C_{4,2}$ are both positive constants independent of Δt.

Proof.
As in the proofs of the previous lemmas, firstly notice that

$$\mathbb{E}\Big[\big\|IV\big\|_2^2\Big] \leqslant IV_1 + IV_2 \tag{2.12}$$

where

$$IV_1 := 2\mathbb{E}\Big[\big\|\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A} DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k)\Delta t d\mathbf{W}(s)\big\|_2^2\Big]$$

and

$$IV_2 := 2\mathbb{E}\Big[\big\|\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A} \Big(\int_0^1 (1-\theta)\mathbf{h}^\top D^2G(\mathbf{u}_k + \theta\mathbf{h})\mathbf{h} d\theta\Big) d\mathbf{W}(s)\big\|_2^2\Big]$$

As regards term IV_1, by Itô's isometry and Hölder's inequality we get

$$IV_1 = 2\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|e^{(t_n-t_{k+\frac{1}{2}})A}DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k)\Delta t\big\|_2^2\Big]ds$$

$$\leqslant 2\sum_{k=0}^{n-1}\Delta t^2 \sup_{0\leqslant k\leqslant n-1} \mathbb{E}\big[\|DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k))\|_F^2\big]\int_{t_k}^{t_{k+1}}\mathbb{E}\big[\|e^{(t_n-t_{k+\frac{1}{2}})A}\|_F^2\big]ds.$$

By Assumptions 1.1.9 and 1.1.10 on $\mathbf{F}$, Assumption 2.1.2 and the boundedness of the exponential operator, we have

$$IV_1 \leqslant C_{4,1}\Delta t^3, \tag{2.13}$$

where $C_{4,1} = \sup\limits_{0\leqslant k\leqslant n-1} \mathbb{E}\big[\|DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k))\|_F^2\big] \sup\limits_{0\leqslant k\leqslant n-1}\mathbb{E}\big[\|e^{(T-t_{k+\frac{1}{2}})A}\|_F^2\big]ds.$

We can now seek for an upper bound for the term IV_2. Since Assumption 2.1.1 holds, we have that $(1-\theta)\mathbf{h}^\top D^2G(\mathbf{u}_k+\theta\mathbf{h})\mathbf{h}$ is bounded in the interval $[0,1]$. Thus,

$$\int_0^1 (1-\theta)\mathbf{h}^\top D^2G(\mathbf{u}_k+\theta\mathbf{h})\mathbf{h}\,d\theta \leqslant \Delta t\max_{0\leqslant\theta\leqslant 1}\Big((1-\theta)\mathbf{h}^\top D^2G(\mathbf{u}_k+\theta\mathbf{h})\mathbf{h}\Big).$$

By Itô's isometry, we further have that

$$IV_2 \leqslant 2\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|e^{(t_n-t_{k+\frac{1}{2}})A}\Delta t\max_{0\leqslant\theta\leqslant 1}\Big((1-\theta)\mathbf{h}^\top D^2G(\mathbf{u}_k+\theta\mathbf{h})\mathbf{h}\Big)\big\|_F^2\Big]ds.$$

By boundedness of the exponential operator

$$IV_2 \leqslant C\Delta t^2\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|\max_{0\leqslant\theta\leqslant 1}\Big((1-\theta)\mathbf{h}^\top D^2G(\mathbf{u}_k+\theta\mathbf{h})\mathbf{h}\Big)\big\|_F^2\Big]ds$$

or

$$IV_2 \leqslant C\Delta t^2\mathbb{E}\Big[\big\|\sum_{k=0}^{n-1}\max_{0\leqslant\theta\leqslant 1}\Big((1-\theta)\mathbf{h}^\top D^2G(\mathbf{u}_k+\theta\mathbf{h})\mathbf{h}\Big)\Delta\mathbf{W}_k\big\|_2^2\Big]. \tag{2.14}$$

Next we notice that, $\mathbf{h}^\top D^2G(\mathbf{u}_k+\theta\mathbf{h})\mathbf{h}\Delta\mathbf{W}_k$ is a vector in $\mathbb{R}^d$ with i-th component given by

$$\sum_{j=1}^{d}\Big(h_1^2\frac{\partial^2 G_{ij}}{\partial u_1^2}+\ldots+h_1h_d\frac{\partial^2 G_{ij}}{\partial u_d\partial u_1}+\ldots+h_dh_1\frac{\partial^2 G_{ij}}{\partial u_1\partial u_d}+\ldots+h_d^2\frac{\partial^2 G_{ij}}{\partial u_d^2}\Big)\Delta W_j$$

$$= \sum_{j,k,\ell=1}^{d} h_kh_\ell\frac{\partial^2 G_{ij}}{\partial u_\ell\partial u_k}\Delta W_j,\quad i=1,\ldots,d, \tag{2.15}$$

where $h_i = F_i(\mathbf{u})\Delta t + \sum_{j=1}^{d} G_{ij}(\mathbf{u})\Delta W_j$. Then $h_k h_\ell$ in (2.15), is given by

$$h_k h_\ell = F_k(\mathbf{u})F_\ell(\mathbf{u})\Delta t^2 + \sum_{j=1}^{d} G_{\ell j}(\mathbf{u})F_k(\mathbf{u})\Delta t\Delta W_j + \sum_{j=1}^{d} G_{kj}(\mathbf{u})G_{\ell j}(\mathbf{u})\Delta W_j\Delta W_j$$
$$+ \sum_{j=1}^{d} G_{kj}(\mathbf{u})F_\ell(\mathbf{u})\Delta t\Delta W_j.$$

So, if we substitute for $h_k h_\ell$ in (2.15), we get that the i-th component of $\mathbf{h}^\top D^2 G(\mathbf{u}_k + \theta\mathbf{h})\mathbf{h}\Delta\mathbf{W}_k$ is given by

$$\sum_{j,k,\ell=1}^{d} \Big(F_k(\mathbf{u})F_\ell(\mathbf{u})\Delta t^2 + \sum_{j=1}^{d} G_{\ell j}(\mathbf{u})F_k(\mathbf{u})\Delta t\Delta W_j + \sum_{j=1}^{d} G_{kj}(\mathbf{u})G_{\ell j}(\mathbf{u})\Delta W_j\Delta W_j$$
$$+ \sum_{j=1}^{d} G_{kj}(\mathbf{u})F_\ell(\mathbf{u})\Delta t\Delta W_j \Big) \frac{\partial^2 G_{ij}}{\partial u_\ell \partial u_k}\Delta W_j.$$

It is then clear to see that if we substitute the above expression in (2.14), the components will be smaller or equal than

$$C\Delta t^2\big(\Delta t^{3/2} + \Delta t^2 + \Delta t^2 + \Delta t^{3/2}\big).$$

This leads to the following bound for the term IV_2

$$IV_2 \leqslant C_{4,2}\big(\Delta t^{7/2} + \Delta t^4\big), \tag{2.16}$$

where $C_{4,2}$ is a positive constant independent of Δt. Putting together (2.13) and (2.16), we get the final estimate. $\square$

Lemma 2.1.7 indicates that the upper bound for the term $\mathbb{E}\big[\|IV\|_2^2\big]$ is significantly smaller than the upper bounds of all the previous terms. This, enables us to ignore the above term when we look for an overall upper bound for the difference $\big\|\mathbf{u}(t_n) - \mathbf{u}_n\big\|_{L^2(\Omega,\mathbb{R}^d)}$. Thus, putting everything from the previous lemmas together, we get that the order of strong convergence for the SEI scheme is given by the following theorem.

Theorem 2.1.8. *Let $\boldsymbol{u}_n$, represented by (2.4), be the approximate solution of the semilinear Stratonovich equation (1.38) at time $t_n = n\Delta t$, $\Delta t > 0$ and $\boldsymbol{u}(t_n)$ represented by (2.1) to be the exact solution of (1.38). Also, let Assumptions 1.1.9,* *1.1.10, 2.1.2 and 2.1.1 hold. Then*

$$\left(\mathbb{E}\left[\left\|\boldsymbol{u}_n - \boldsymbol{u}(t_n)\right\|_2^2\right]\right)^{1/2} \leqslant C\sqrt{\Delta t}|\ln(\Delta t)|,$$

where C is a positive constant independent of Δt.

Proof.
From Lemmas 2.1.3, 2.1.5 and 2.1.6 we have that

$$\mathbb{E}\left[\|\mathbf{u}(t_n) - \mathbf{u}_n\|_2^2\right] \leqslant K\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\left[\|\mathbf{u}(s) - \mathbf{u}_k\|_2^2\right] + L\Delta t\big(\ln(\Delta t)\big)^2,$$

where K, L are positive constants, $K = C_1 + C_{2,1} + C_{3,1}$ and $L = C_{2,2} + C_{3,1}$. By Gronwall's lemma we get

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{u}(t_n) - \mathbf{u}_n\|_2^2\right] &\leqslant L\Delta t\big(\ln(\Delta t)\big)^2 + K\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{K\Delta t(t-s)}\Delta t\big(\ln(\Delta t)\big)^2 ds \\
&= L\Delta t\big(\ln(\Delta t)\big)^2 + K\Delta t\big(\ln(\Delta t)\big)^2\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{K\Delta t(t-s)} ds \\
&\leqslant C\Delta t\big(\ln(\Delta t)\big)^2.
\end{aligned}$$

So, taking the square root gives

$$\left(\mathbb{E}\left[\left\|\mathbf{u}_n - \mathbf{u}(t_n)\right\|_2^2\right]\right)^{1/2} \leqslant C\sqrt{\Delta t}|\ln(\Delta t)|.$$

□

Practically, this means that the order of convergence of the SEI scheme for the general case, where $G(\mathbf{u}) \in \mathbb{R}^{d\times d}$ is the non-diagonal diffusion term, is $O(\Delta t^{1/2-\epsilon})$, for any $\epsilon > 0$. This can also be seen in the numerical examples of the following section.

Before moving to the numerical examples, we should not fail to attract the reader's attention to the fact that although we obtained the predictor $\tilde{\mathbf{u}}_{n+1}$ by the Euler-Maruyama scheme here, it may as well be obtained by other numerical schemes like for example by the Itô exponential integrators scheme

$$\tilde{\mathbf{u}}_{n+1} = e^{\Delta tA}\mathbf{u}_n + A^{-1}(e^{\Delta tA} - I)\mathbf{F}(\mathbf{u}_n) + e^{\Delta tA}G(\mathbf{u}_n)\Delta\mathbf{W}_n.$$

In fact, the proof of convergence itself can be generalised to the case where $\tilde{\mathbf{u}}_{n+1}$ is given by any numerical scheme with known rate of convergence. More details about this generalisation are given later in Chapter 3.

2.2 Strong convergence for SEI numerically

In order to investigate strong convergence for SEI numerically, we consider the difference between the 'exact' and the approximate solution. What is considered as exact solution in this context is denoted by $\mathbf{u}$ref, while the approximate solutions for different step-sizes are denoted by $\mathbf{u}$dt. The solution $\mathbf{u}$ref is given by SEI using a small step-size Dt, considered as the reference step-size, while the solutions $\mathbf{u}$dt are given by the same scheme with bigger step-sizes $k \cdot Dt$, multiples of the reference one. The mean is obtained by considering a sample average. In other words, we generate M sample paths given from the SEI scheme for k different step-sizes and we take the average difference between $\mathbf{u}\text{ref}_n$ and $\mathbf{u}\text{dt}_n$, where n is the final time step.

Algorithm 7.3.7 generates the sample paths from the SEI scheme. The inputs are the initial condition $\mathbf{u}0$, the time interval $[0, T]$, the number of iterations N, the drift and diffusion terms $A\mathbf{u} + \mathbf{F}(\mathbf{u})$ and $G(\mathbf{u})$ respectively and their dimensions d and m. Other inputs are the vector *kappa*, that consists of integer elements which are multiplied by Dt so as to create the different step-sizes as well as M, which as mentioned previously denotes the number of the generated samples. The outputs are the solution $\mathbf{u}$ and the time vector t. Note that the random variable $\mathbf{u}$ is a three dimensional array where $\mathbf{u}(\ell, j, n)$, $\ell = 1, \ldots, d$, $j = 1, \ldots, M$, and $n = 1, \ldots, N+1$.

Algorithm 7.3.8 plots the error between the 'exact' and the approximate solutions. This is achieved by considering the difference between $\mathbf{u}$ref and $\mathbf{u}$dt for all the different step-sizes and for all the samples, taking the norm of the differences and the average of the norms.

We consider examples of semilinear Stratonovich SDEs in one, two and three dimensions and we examine both diagonal and non-diagonal multiplicative noise. In case of diagonal noise, $G(\mathbf{u})$ is a $d \times d$ matrix with non-zero coefficients in the diagonal and zeros everywhere else, while in case of non-diagonal noise $G(\mathbf{u})$ is a $d \times d$ matrix with non-zero coefficients. In Figure 2.1 we plot the solution of a toy example in one dimension while in Figures 2.2, 2.3, 2.5 and 2.6 we show convergence for different examples in two and three dimensions. Later, we use SEI to approximate solutions for the Landau-Lifshitz-Gilbert equation. We notice that strong convergence is of order $O(\Delta t)$ for diagonal noise and of order $O(\sqrt{\Delta t})$ for non-diagonal noise.

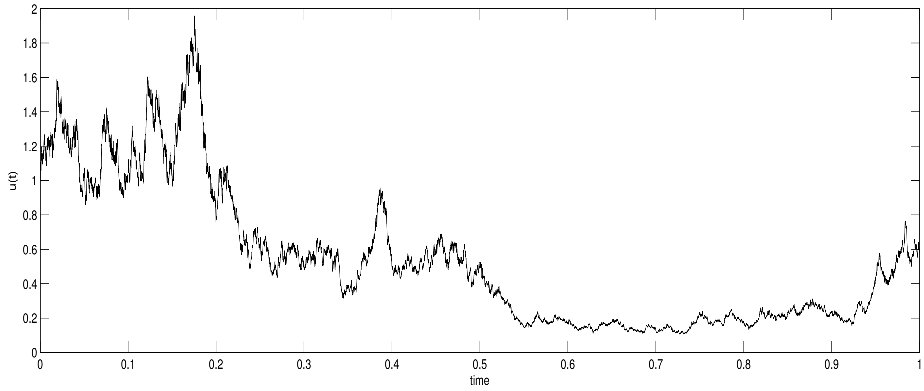


Figure 2.1: Approximate solution of $du = \big(Au + f(u)\big)dt + g(u) \circ dW$ with $f(u) = u$, $g(u) = -2u$, $A = 1$, after 10^4 iterations and with initial condition $u_0 = 1$.

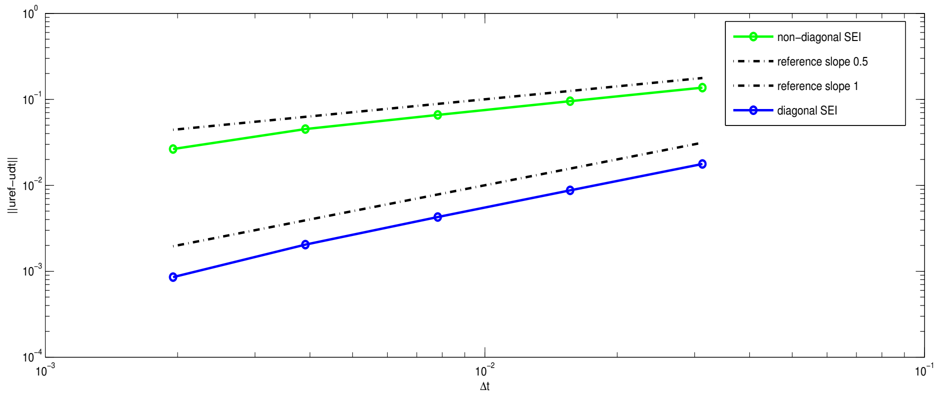


Figure 2.2: Green line shows strong convergence of order $O(\sqrt{\Delta t})$ for 2-dimensional system with non-diagonal noise. More specifically, $\mathbf{F}(\mathbf{u}) = 0.2\mathbf{u}$, $G(\mathbf{u}) = \begin{pmatrix} 0.2u_1 & 2 \\ 2 & 0.2u_2 \end{pmatrix}$. Bottom blue line shows strong convergence of order $O(\Delta t)$ for 2-dimensional system with diagonal noise. Drift term is the same and diffusion is given by $G(\mathbf{u}) = \begin{pmatrix} 0.2u_1 & 0 \\ 0 & 0.2u_2 \end{pmatrix}$. $M = 100$ samples, reference step-size $\frac{T}{N} = 2^{-11}$.

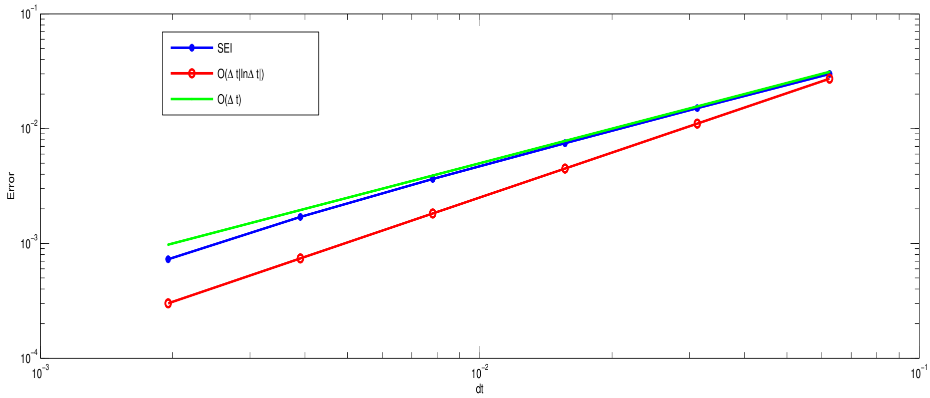


Figure 2.3: Diagonal noise, three-dimensional example, that is $\mathbf{u} \in \mathbb{R}^3$ and $G(\mathbf{u}) \in \mathbb{R}^{3\times 3}$ is a 3×3 matrix. Strong order of convergence is $O(\Delta t)$ for this example with $M = 300$ samples and $N = 2^{11}$.

Another example that we examine numerically is the stochastic Landau-Lifshitz-Gilbert (LLG) equation. The LLG equation is commonly used to model the evolution of magnetisation in a solid and more particularly in a ferromagnetic material.We look at the stochastic version of the LLG model for which there is evidence that it should be interpreted in the Stratonovich sense, see [10] and references therein.

On a bounded domain $D \subset \mathbb{R}^2$ let $D_T := (0,T) \times D$, $\mathbf{M} : D_T \to \mathbb{S}^2 := \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| = 1\}$ be the magnetisation and let $\mathbf{h} : D_T \to \mathbb{R}^3$ denote the effective field. The stochastic LLG is interpreted in the Stratonovich sense and is given by

$$d\mathbf{M}(t,\mathbf{x}) = \mathbf{M}(t,\mathbf{x}) \times \big[\mathbf{h}(t,\mathbf{x}) - \alpha\mathbf{M}(t,\mathbf{x}) \times \mathbf{h}(t,\mathbf{x})\big]dt + \sigma\mathbf{M}(t,\mathbf{x}) \times \circ d\mathbf{W}(t,\mathbf{x}), \quad (2.17)$$

where $\alpha > 0$ is a damping parameter, σ is a constant that determines the noise intensity and $\mathbf{W}$ is an $\mathbb{R}^3$-valued Wiener process. Following the same order with [5], we start from studying the Langevin dynamics for a finite number of ferromagnetic spins N. In other words, we start from the system of SDEs that corresponds to the SPDE (2.17) as given in the following equation.

$$d\mathbf{m}_i(t) = \mathbf{m}_i(t) \times \big[\mathbf{h}_i(\mathbf{m}(t)) - \alpha\mathbf{m}_i(t) \times \mathbf{h}_i(\mathbf{u}(t))\big]dt + \sigma\mathbf{m}_i(t) \times \circ d\mathbf{W}_i(t), \quad (2.18)$$

where $i = 1, \ldots, N$, $\mathbf{m} \in (\mathbb{R}^3)^N$ with $\mathbf{m} = (\mathbf{m}_1, \ldots, \mathbf{m}_N)^\top$, $\mathbf{W}_i$ are $\mathbb{R}^3$-valued Wiener processes and we consider some initial condition $\mathbf{m}_i(0)$ that satisfies $|\mathbf{m}_i(0)| = 1$.

More specifically, we focus on the single spin example, i.e. $N = 1$ and we switch the notation to $\mathbf{u}(t)$ instead of $\mathbf{m}(t)$ in order to be consistent with the notation that we use throughout the thesis. Hence, (2.18) for $N = 1$ reads

$$d\mathbf{u}(t) = \mathbf{u}(t) \times \big[\mathbf{h}(\mathbf{u}(t)) - \alpha\mathbf{u}(t) \times \mathbf{h}(\mathbf{u}(t))\big]dt + \sigma\mathbf{u}(t) \times \circ d\mathbf{W}(t), \quad (2.19)$$

with $\mathbf{u} = (u_1, u_2, u_3)^\top$, initial condition $\mathbf{u}(0) = \mathbf{u}_0 = (0, 1, 0)^\top$ and α, σ being the regulating parameters as before. By using the distributivity rule of the cross product, we express the three-dimensional system in a component fashion.

$$
\begin{aligned}
du_1 &= (\alpha u_1^2 + \alpha u_3^2)dt + \sigma u_2 \circ dW_3 - \sigma u_3 \circ dW_2 \qquad (2.20)\\
du_2 &= (u_3 - \alpha u_1 u_2)dt + \sigma u_3 \circ dW_1 - \sigma u_1 \circ dW_2\\
du_3 &= (-\alpha u_1 u_3 - u_2)dt + \sigma u_1 \circ dW_2 - \sigma u_2 \circ dW_1.
\end{aligned}
$$

We approximate the solution of the system (2.20) by using both SEI, as given in equation (1.43) and the Euler-Maruyama scheme. Next, we compare the two results in Figure 2.4. Note that in order to use Euler-Maruyama, it is required to write (2.20) as its equivalent Itô-interpreted system according to the Itô-Stratonovich correction formula, more details of this can be found in Example 5 of the Appendix Chapter 7. As expected from our analysis, the convergence of SEI scheme is approximately of order $O(\sqrt{\Delta t})$, more precisely of order $O(\Delta t^{1/2-\epsilon})$ for some small $\epsilon > 0$, since we have non-diagonal noise. The convergence rate for the non-diagonal noise case can be seen in Figure 2.5 while in Figure 2.6, we solve a similar to (2.20) system with diagonal noise.

Our numerical simulations indicate that unlike the approach of [5], our scheme does not guarantee the conservation of magnitude of magnetisation i.e.$|\mathbf{m}(t)| = |\mathbf{m}(0)| = 1$. Of course one could overcome this by considering some type of artificial normalisation. In that case though, convergence needs to be reexamined.

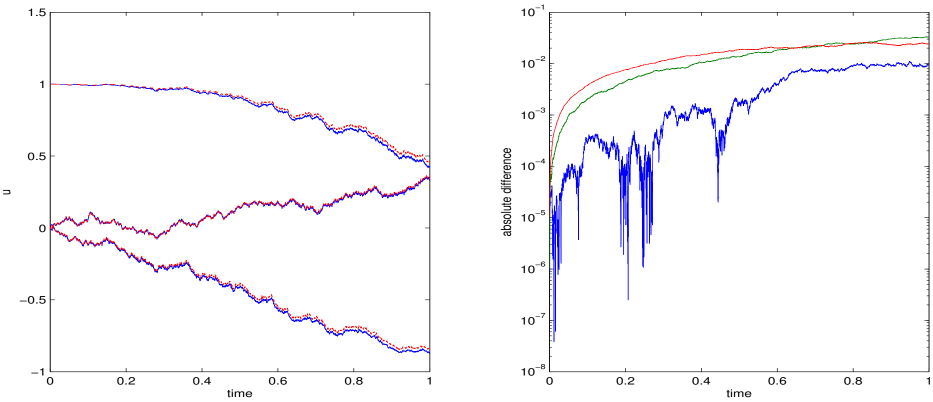


Figure 2.4: LHS:Three-dimensional Langevin system approximated by SEI in blue and by corrected Euler-Murayama in red. RHS: Absolute difference between the two approximations against time, in a log-log scale. Initial condition $\mathbf{u}_0 = (0, 1, 0)^\top$, noise intensity $\sigma = 0.2$, damping parameter $\alpha = 0.25$.

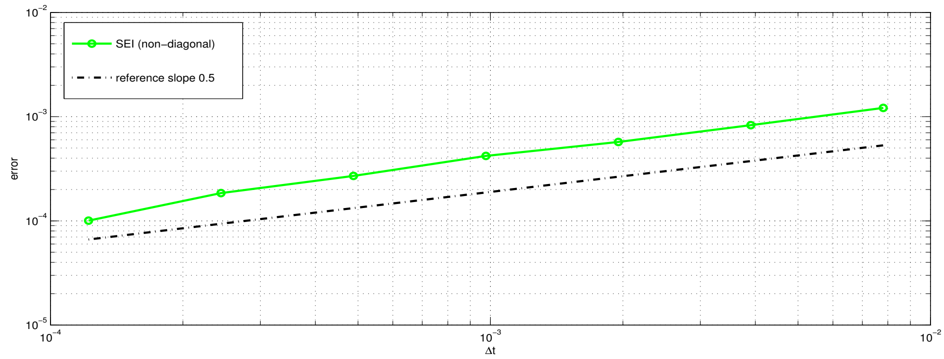


Figure 2.5: Strong convergence for SEI when solving the three-dimensional Langevin system (2.20). Using polyfit we get that the slope of the fitted line is 0.578590.

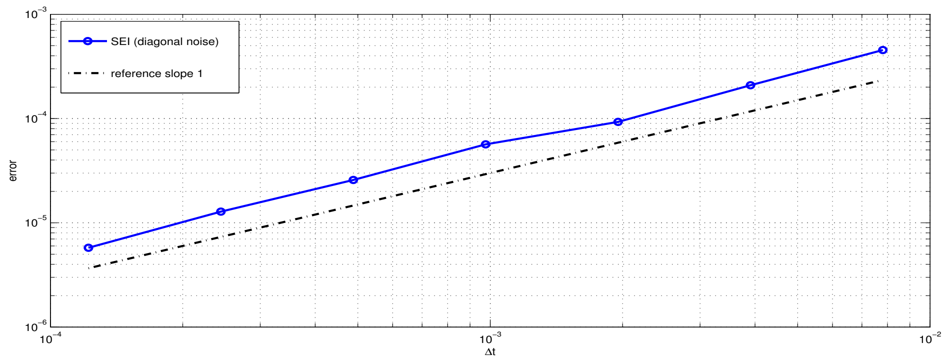


Figure 2.6: strong convergence for SEI solving a three-dimensional system of SDEs similar to (2.20) but with diagonal noise. Diffusion term is given by `ghandle=@(u1,u2,u3) [-sigma*u1 0 0;0 -sigma*u2 0;0 0 -sigma*u3]`. The slope of the fitted line is 1.028944.

2.3 Higher order exponential Milstein-type scheme for Stratonovich SDEs

In a similar way with the exponential Milstein-type Itô scheme of Section 1.2.4, in this section we derive a new higher order scheme for semilinear Stratonovich SDEs of the form (1.38).

2.3.1 MSEI2

Under Assumptions 1.1.9, 1.1.10 and 2.1.2, the mild solution of equation (1.38) considered in the interval $[t, r]$, satisfies

$$\mathbf{u}(t) = e^{(t-r)A}\mathbf{u}(r) + \int_r^t e^{(t-s)A}\mathbf{F}(\mathbf{u}(s))ds + \int_r^t e^{(t-s)A}G(\mathbf{u}(s)) \circ d\mathbf{W}(s). \qquad (2.21)$$

We Taylor expand $\mathbf{F}(\mathbf{u}(s))$ about the left end point $\mathbf{u}(r)$ and $G(\mathbf{u}(s))$ about the midpoint $\mathbf{u}\big(\frac{r+t}{2}\big)$. However, instead of writing $G\big(\mathbf{u}\big(\frac{r+t}{2}\big)\big)$, we further approximate to $\frac{1}{2}\big(G(\mathbf{u}(r))+G(\mathbf{u}(t))\big)$ since this appears to be a natural choice for the approximation of the Stratonovich integral. Moreover, for the same reason, instead of $\mathbf{u}\big(\frac{r+t}{2}\big)$ we look at $\frac{1}{2}\big(\mathbf{u}(r)+\mathbf{u}(t)\big)$. By Taylor expanding $\mathbf{u}\big(\frac{r+t}{2}\big)$ twice, once about $\mathbf{u}(r)$ and once about $\mathbf{u}(t)$, it is easy to check that $\mathbf{u}\big(\frac{r+t}{2}\big) \approx \frac{1}{2}\big(\mathbf{u}(r)+\mathbf{u}(t)\big)$ when $[t,r]$ is small. Then, the solution satisfies

$$\mathbf{u}(t) \simeq e^{(t-r)A}\mathbf{u}(r) + \int_r^t e^{(t-s)A}\mathbf{F}(\mathbf{u}(r))ds + \frac{1}{2}\int_r^t e^{(t-s)A}\big(G(\mathbf{u}(r)) + G(\mathbf{u}(t))\big) \circ d\mathbf{W}(s)$$
$$+ \int_r^t e^{(t-s)A}\Big(\mathbf{u}(s) - \frac{\mathbf{u}(r)+\mathbf{u}(t)}{2}\Big)\frac{DG(\mathbf{u}(r)) + DG(\mathbf{u}(t))}{2} \circ d\mathbf{W}(s).$$

Next, we substitute for the difference $\Big(\mathbf{u}(s) - \frac{\mathbf{u}(r)+\mathbf{u}(t)}{2}\Big)$ by the Euler-Maruyama scheme. Recall from the Milstein and exponential Milstein schemes for Itô SDEs, that when we substitute the above difference by the Euler-Maruyama approximation, we omit the term that corresponds to the approximation of the drift, see Subsections 1.2.2 and 1.2.4. Hence, we get

$$\mathbf{u}(t) \simeq e^{(t-r)A}\mathbf{u}(r) + \int_r^t e^{(t-s)A}\mathbf{F}(\mathbf{u}(r))ds + \frac{1}{2}\int_r^t e^{(t-s)A}\big(G(\mathbf{u}(r)) + G(\mathbf{u}(t))\big) \circ d\mathbf{W}(s)$$
$$+ \frac{DG(\mathbf{u}(r)) + DG(\mathbf{u}(t))}{2}\int_r^t e^{(t-s)A}\left(\frac{G(\mathbf{u}(r)) + G(\mathbf{u}(t))}{2}\int_{\frac{r+t}{2}}^s d\mathbf{W}(\tau)\right) \circ d\mathbf{W}(s). \tag{2.22}$$

We expand the double stochastic integral of (2.22), for details see Examples 6 and 7 of the Appendix, to get

$$\frac{DG(\mathbf{u}(r)) + DG(\mathbf{u}(t))}{2}\int_r^t e^{(t-s)A}\left(\frac{G(\mathbf{u}(r)) + G(\mathbf{u}(t))}{2}\int_{\frac{r+t}{2}}^s d\mathbf{W}(\tau)\right) \circ d\mathbf{W}(s) =$$
$$\frac{1}{8}e^{\frac{(t-r)A}{2}}\big(DG(\mathbf{u}(r)) + DG(\mathbf{u}(t))\big)\big(G(\mathbf{u}(r)) + G(\mathbf{u}(t))\big)\Big[\big(\mathbf{W}(t) - \mathbf{W}(\frac{t+r}{2})\big)^2$$
$$-\big(\mathbf{W}(r) - \mathbf{W}(\frac{t+r}{2})\big)^2\Big].$$

The above integral considered in the time interval $[t_n, t_{n+1}]$ reads

$$\frac{DG(\mathbf{u}(t_n)) + DG(\tilde{\mathbf{u}}(t_{n+1}))}{2}\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\frac{G(\mathbf{u}(t_n)) + G(\tilde{\mathbf{u}}(t_{n+1}))}{2}\int_{t_{n+\frac{1}{2}}}^s d\mathbf{W}(\tau) \circ d\mathbf{W}(s)$$

$$= \frac{1}{8} e^{\frac{\Delta t A}{2}} \big(DG(\mathbf{u}(t_n)) + DG(\tilde{\mathbf{u}}(t_{n+1})) \big) \big(G(\mathbf{u}(t_n)) + G(\tilde{\mathbf{u}}(t_{n+1})) \big)$$

$$\times \Big[\big(\mathbf{W}(t_{n+1}) - \mathbf{W}(\frac{t_{n+1}+t_n}{2}) \big)^2 - \big(\mathbf{W}(t_{n+1}) - \mathbf{W}(\frac{t_{n+1}+t_n}{2}) \big)^2 \Big],$$

where $\tilde{\mathbf{u}}_{n+1}$ is taken by the scheme of Section 1.2.3 as it is given in equation (1.24).

This leads to a higher order scheme, given in component form by the following equation. We call this scheme Milstein-type Stratonovich Exponential Integrators 2 (MSEI2) and as we did with SEI, we use it to approximate solutions of the three-dimensional Stratonovich Landau-Lifshitz-Gilbert equation (2.20). Our numerical experiments indicate that MSEI2 converges with a strong order of approximately $O(\Delta t)$ for the case of multiplicative noise, as seen in Figure 2.7.

$$u_{k,n+1} = \sum_{q=1}^{d} a_{kq} u_q + \sum_{q=1}^{d} a_{kq} \mathbf{F}_q(\mathbf{u}_n) + \sum_{q=1}^{d} \sum_{j=1}^{m} \hat{a}_{kq} \big(G_{qj}(\mathbf{u}_n) + G_{qj}(\tilde{\mathbf{u}}_{n+1}) \big) \Delta W_{j,n} \tag{2.23}$$

$$+ \frac{1}{8} \sum_{q=1}^{d} \sum_{j=1}^{m} \sum_{\ell=1}^{d} \hat{a}_{kq} \Big(\frac{\partial G_{qj}}{\partial u_\ell}(\mathbf{u}_n) G_{\ell j}(\mathbf{u}_n) + \frac{\partial G_{qj}}{\partial u_\ell}(\tilde{\mathbf{u}}_{n+1}) G_{\ell j}(\tilde{\mathbf{u}}_{n+1}) \Big) I_j(t_{n+\frac{1}{2}}, s) J_j(t_n, t_{n+1})$$

$$+ \frac{1}{8} \sum_{q=1}^{d} \sum_{j<i=1}^{m} \sum_{\ell=1}^{d} \hat{a}_{kq} \Bigg[\Bigg(\frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n) G_{\ell j}(\mathbf{u}_n) + \frac{\partial G_{qi}}{\partial u_\ell}(\tilde{\mathbf{u}}_{n+1}) G_{\ell j}(\tilde{\mathbf{u}}_{n+1}) \Bigg)$$

$$+ \Bigg(\frac{\partial G_{qj}}{\partial u_\ell}(\mathbf{u}_n) G_{\ell i}(\mathbf{u}_n) + \frac{\partial G_{qj}}{\partial u_\ell}(\tilde{\mathbf{u}}_{n+1}) G_{\ell i}(\tilde{\mathbf{u}}_{n+1}) \Bigg) \Bigg] I_i(t_{n+\frac{1}{2}}, s) J_j(t_n, t_{n+1})$$

$$+ \frac{1}{8} \sum_{q=1}^{d} \sum_{j<i=1}^{m} \sum_{\ell=1}^{d} \hat{a}_{kq} \Bigg[\Bigg(\frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n) G_{\ell j}(\mathbf{u}_n) + \frac{\partial G_{qi}}{\partial u_\ell}(\tilde{\mathbf{u}}_{n+1}) G_{\ell j}(\tilde{\mathbf{u}}_{n+1}) \Bigg)$$

$$+ \Bigg(\frac{\partial G_{qj}}{\partial u_\ell}(\mathbf{u}_n) G_{\ell i}(\mathbf{u}_n) + \frac{\partial G_{qj}}{\partial u_\ell}(\tilde{\mathbf{u}}_{n+1}) G_{\ell i}(\tilde{\mathbf{u}}_{n+1}) \Bigg) \Bigg] A_{ij,n},$$

where a_{kq}, $\hat{a}_{kq}$ are the kq-elements of the operators $e^{\Delta t A}$, $e^{\frac{\Delta t A}{2}}$ respectively. Also, $I_i(\frac{t_n+t_{n+1}}{2}, s) = I_i(t_{n+\frac{1}{2}}, s) := \int_{t_{n+\frac{1}{2}}}^{s} dW_i(\tau)$ and $J_j(t_n, t_{n+1}) := \int_{t_n}^{t_{n+1}} \circ dW_j(s)$.

2.3.2 MSEI

An alternative way for deriving a similar higher order scheme is to change slightly our approximation by using a combination of the left-end point and the mid-point for evaluating $G(\mathbf{u}(s))$. This way, may be seen as an extra approximating step, where we Taylor expand $\mathbf{u}(t)$ about $\mathbf{u}(r)$ so that $\mathbf{u}(t) = \mathbf{u}(r) + D\mathbf{u}(r)(t-r) + h.o.t.$ Then, $\dfrac{\mathbf{u}(t)+\mathbf{u}(r)}{2} = \mathbf{u}(r) + D\mathbf{u}(r)(t-r) + h.o.t.$ Note that, in the discretised

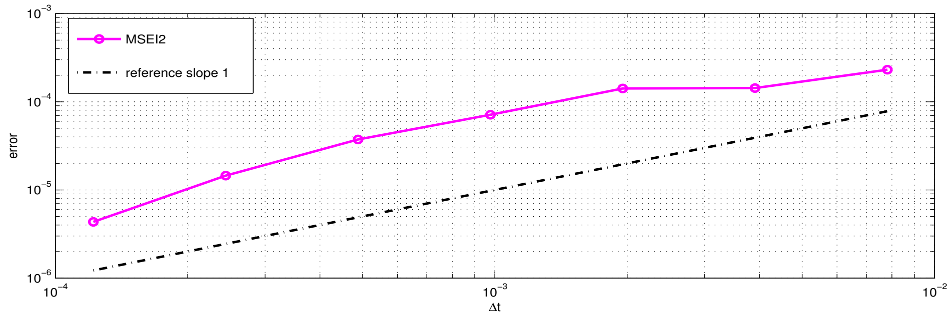


Figure 2.7: Strong convergence for MSEI 2 scheme approximating the LLG system (2.20) with non diagonal, commutative noise against reference line with slope 1. M=100 samples, reference time step-size $\Delta t = 2^{-14}$. The slope of the fitted line is 0.918767.

version that we consider later on, the interval $(t - r)$ becomes $(t_{n+1} - t_n)$, that is $\dfrac{\mathbf{u}(t) + \mathbf{u}(r)}{2} \approx \mathbf{u}(r) + O(\Delta t)$.

So, looking back at equation (2.21), we substitute for $\mathbf{F}(\mathbf{u}(s))$, $G(\mathbf{u}(s))$ and for the difference $\mathbf{u}(s) - \mathbf{u}(r)$ given by the Euler-Maruyama scheme, skipping the drift term as before. Hence, our approximation becomes

$$\mathbf{u}(t) = e^{(t-r)}\mathbf{u}(r) + \int_r^t e^{(t-s)A}\mathbf{F}(\mathbf{u}(r))ds + \frac{1}{2}\int_r^t e^{(t-s)A}\big(G(\mathbf{u}(r)) + G(\mathbf{u}(t))\big) \circ d\mathbf{W}(s)$$

$$+ \int_r^t e^{(t-s)A}DG(\mathbf{u}(r))\Big(G(\mathbf{u}(r))\int_r^s d\mathbf{W}(\tau)\Big) \circ d\mathbf{W}(s) + R_M(s; r, \mathbf{u}(r)), \qquad (2.24)$$

where R_M is the remainder term from the expansions of $\mathbf{F}$ and G and the error from using Euler-Maruyama denoted by R_E

$$R_M(s; r, \mathbf{u}(r)) = \int_r^t e^{(t-s)A}DF(\mathbf{u}(r))\big(\mathbf{u}(s) - \mathbf{u}(r)\big)ds$$

$$+ \int_r^t e^{(t-s)A}\int_0^1 (1-\theta)D^2G\big(\mathbf{u}(r) + \theta(\mathbf{u}(s)) - \mathbf{u}(r)\big)(\mathbf{u}(s) - \mathbf{u}(r))^2 d\theta \circ d\mathbf{W}(s)$$

$$\int_r^t e^{(t-s)A}DG(\mathbf{u}(r))\big(\mathbf{F}(\mathbf{u}(r))(t-r) + R_E\big) \circ d\mathbf{W}(s). \qquad (2.25)$$

Then, equation (2.24) considered in the time interval $[t_n, t_{n+1}]$ yields

$$\mathbf{u}(t_{n+1}) = e^{\Delta tA}\mathbf{u}(t_n) + \mathbf{F}(\mathbf{u}(t_n))\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}ds$$

$$+\frac{1}{2}\int_{t_n}^{t_{n+1}} e^{\frac{\Delta tA}{2}}\big(G(\mathbf{u}(t_n)) + G(\tilde{\mathbf{u}}(t_{n+1}))\big)d\mathbf{W}(s)$$

$$+ \int_{t_n}^{t_{n+1}} e^{\frac{\Delta tA}{2}}DG(\mathbf{u}(t_n))\Big(G(\mathbf{u}(t_n))\int_{t_n}^s d\mathbf{W}(\tau)\Big) \circ d\mathbf{W}(s) + R_M(s; t_n, \mathbf{u}(t_n)). \quad (2.26)$$

After expanding the double integral term in (2.26) we obtain the k-th component of the scheme that approximates solutions to equation (1.38) for general, i.e. non-diagonal, non-commutative, multiplicative noise.

$$u_{k,n+1} = \sum_{q=1}^{d} a_{kq} u_q + \sum_{q=1}^{d} a_{kq} \mathbf{F}_q(\mathbf{u}_n) + \sum_{q=1}^{d}\sum_{j=1}^{m} \hat{a}_{kq}\big(G_{qj}(\mathbf{u}_n) + G_{gj}(\tilde{\mathbf{u}}_{n+1})\big)\Delta W_{j,n} \tag{2.27}$$

$$+ \frac{1}{2}\sum_{q=1}^{d}\sum_{i=1}^{m}\sum_{\ell=1}^{d} \hat{a}_{kq}\frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell i}(\mathbf{u}_n)\Delta W_{j,n}^2$$

$$+ \frac{1}{2}\sum_{q=1}^{d}\sum_{i<j=1}^{m}\sum_{\ell=1}^{d} \hat{a}_{kq}\Big(\frac{\partial G_{qj}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell i}(\mathbf{u}_n) + \frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell j}(\mathbf{u}_n)\Big) I_i(t_n, s)J_j(t_n, t_{n+1})$$

$$+ \frac{1}{2}\sum_{q=1}^{d}\sum_{i<j=1}^{m}\sum_{\ell=1}^{d} \hat{a}_{kq}\Big(\frac{\partial G_{qj}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell i}(\mathbf{u}_n) + \frac{\partial G_{qi}}{\partial u_\ell}(\mathbf{u}_n)G_{\ell j}(\mathbf{u}_n)\Big) A_{ij,n},$$

where a_{kq} and $\hat{a}_{kq}$, as before are the kq-elements of the operators $e^{\Delta t A}$ and $e^{\frac{\Delta t A}{2}}$ respectively and $I_i(t_n, s) = \displaystyle\int_{t_n}^{s} dW(\tau)$, $J_j(t_n, t_{n+1}) = \displaystyle\int_{t_n}^{t_{n+1}} \circ dW(s)$.

We call this scheme Milstein-type Stratonovich Exponential Integrators (MSEI) and its component wise form (2.27) is the one to be used for numerical implementation. We use MSEI to approximate solutions of the Stratonovich LLG equation in Figure 2.8 where we compare it with the approximation given by the corrected Milstein scheme. In Figures 2.9 and 2.10, we compare our two approaches MSEI2 and MSEI and by computing the absolute difference between the different approximations, we deduce that they agree.

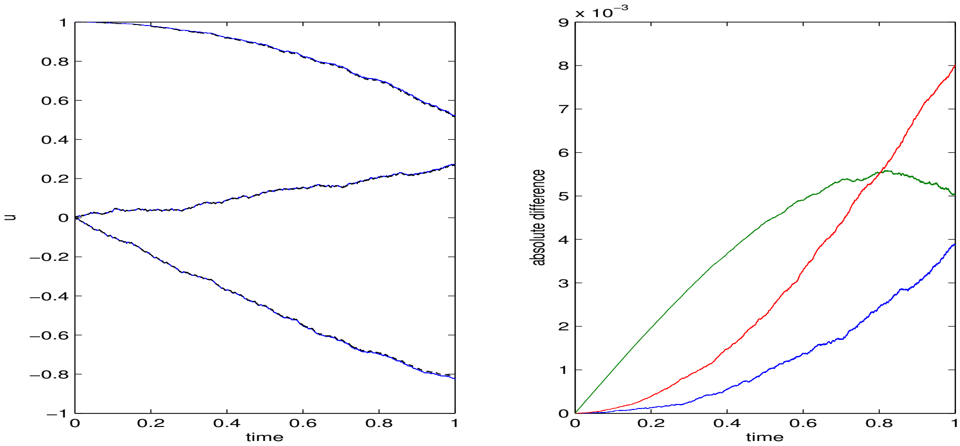


Figure 2.8: Comparison between MSEI in black dashed line and Milstein in blue solid line, both for Stratonovich interpreted 3-d Langevin system. Noise intensity $\sigma = 0.1$.

Note that $\tilde{\mathbf{u}}_{n+1}$ is as before a predictor that may be taken by Euler-Maruyama or other exponential scheme. Alternatively, instead of approximating $\tilde{\mathbf{u}}_{n+1}$, one can expand $G(\tilde{\mathbf{u}}_{n+1})$ around the point $\mathbf{u}(t_n)$ for example. We follow the latter approach together with a continuous version of the scheme that is more useful for our analysis later. We compare the two different approaches in Figure 2.14.

In case of diagonal or commutative noise, recall from Subsection 1.2.4 that some reductions occur. These reductions apparently affect the order of convergence for the scheme. More specifically, the order of convergence for MSEI in case of non-diagonal commutative noise is approximately $O(\Delta t)$, as seen in Figure 2.12. As expected, this is slightly faster compared to the general non-commutative noise case where the order is approximately $O(\Delta t^{0.8})$, see Figure 2.13 and compare Figure 2.11 with Figure 2.12. The small reduction in the order of convergence is explained by the fact that in case of non-commutative noise, our scheme requires the approximation of the double integral term $A_{ij,n}$, which is achieved by solving the auxiliary system (1.20) by means of the Euler-Maruyama method.

In fact, there are two different step-sizes involved in our approximation, one for the Euler-Maruyama approximation of the auxiliary system and one for the MSEI scheme itself. Apparently, as Figure 2.11 suggests, using the same small step-size between Euler-Maruyama and MSEI, could help in achieving an overall order of strong convergence better than $O(\Delta t^{0.8})$ for the general case. This result is later confirmed by our analysis, where be prove strong convergence of order one for the general case.

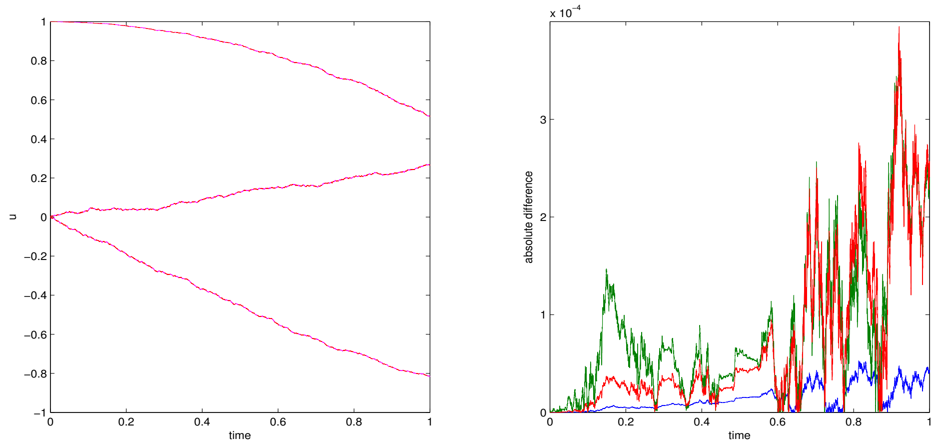


Figure 2.9: On the left, comparison between MSEI2 (magenta solid line) and MSEI (red dashed line) for approximating the Stratonovich Langevin system. On the right, absolute difference between the two different approximations against time.

The performances of MSEI2 and MSEI are similar, yet the computer process time for the second is slightly smaller. This is true both in the commutative and the non-commutative noise cases, as seen in Figures 2.12 and 2.13 respectively. Both MSEI2 and MSEI achieve a higher order of convergence for the non-diagonal noise case compared to the SEI scheme, analysed in Section 2.1.

Another factor that slightly affects the performance of MSEI is the use of predictor for the value of $\tilde{\mathbf{u}}_{n+1}$. In our numerical approximations so far, we obtained $\tilde{\mathbf{u}}_{n+1}$ by means of the exponential scheme of Subsection 1.2.3, we see later in Chapter 3 however, that we can substitute $\tilde{\mathbf{u}}_{n+1}$ by any numerical scheme. In Figure 2.14 we compare the convergence rate and the computer process times of the case where we obtain $\tilde{\mathbf{u}}_{n+1}$ by the scheme of Section 1.2.3 against the case where we just approximate $G(\tilde{\mathbf{u}}_{n+1})$ by means of a Taylor expansion. As expected, convergence is slightly faster when using a predictor, whereas we get slightly bigger computer process times.

Before moving on to the analysis, we mention another approach that might worth considering in the future. Looking back at the derivation of MSEI scheme and in particular at equation (2.24), instead of using Euler-Maruyama to obtain the difference $\mathbf{u}(s) - \mathbf{u}(r)$, we could use the ETD scheme. Since the order of convergence for the two schemes is approximately the same, the above substitution is not expected to affect the overall order of convergence for MSEI. It might however increase the efficiency of the scheme.

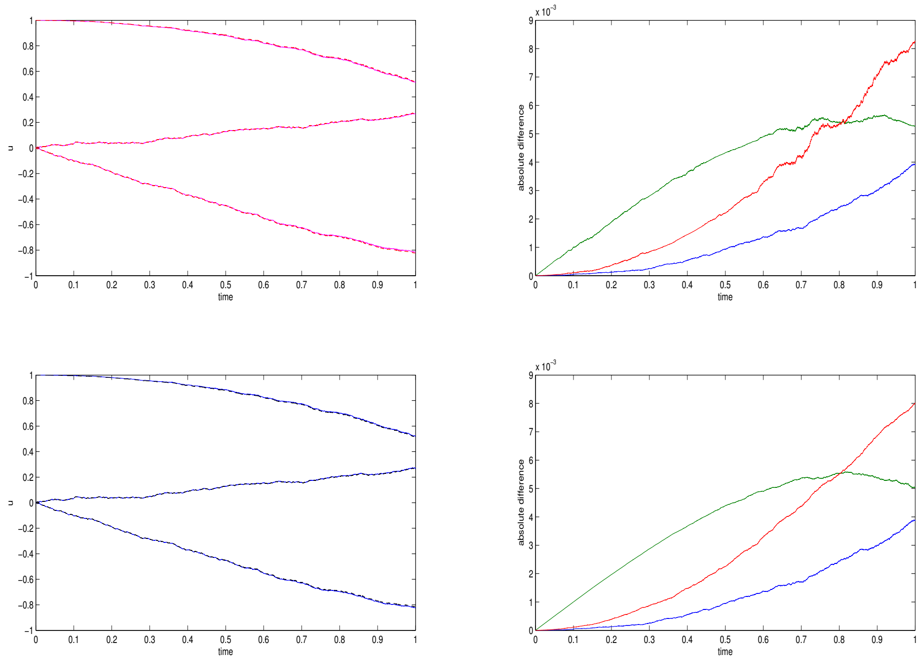


Figure 2.10: Top figures: comparison between MSEI2 (magenta solid line) and corrected exponential Milstein (red solid line), bottom figures: comparison between MSEI (black dashed line) and corrected exponential Milstein (blue solid line), both for approximating Stratonovich Langevin system. Noise intensity $\sigma = 0.1$.

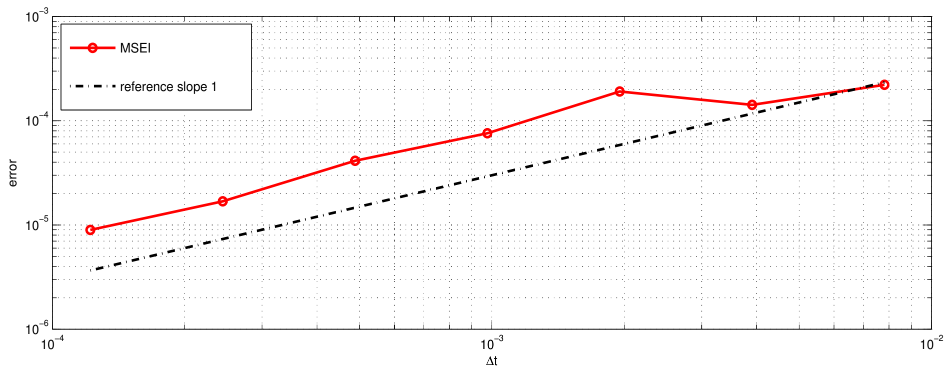


Figure 2.11: Stratonovich exponential Milstein (MSEI) for 3-d LLG (2.20) with non-diagonal, non-commutative, multiplicative noise without using a predictor. $M = 100$ samples, noise intensity $\sigma = 0.1$, reference time step-size 2^{-14}. The slope of the fitted line is 0.794216.

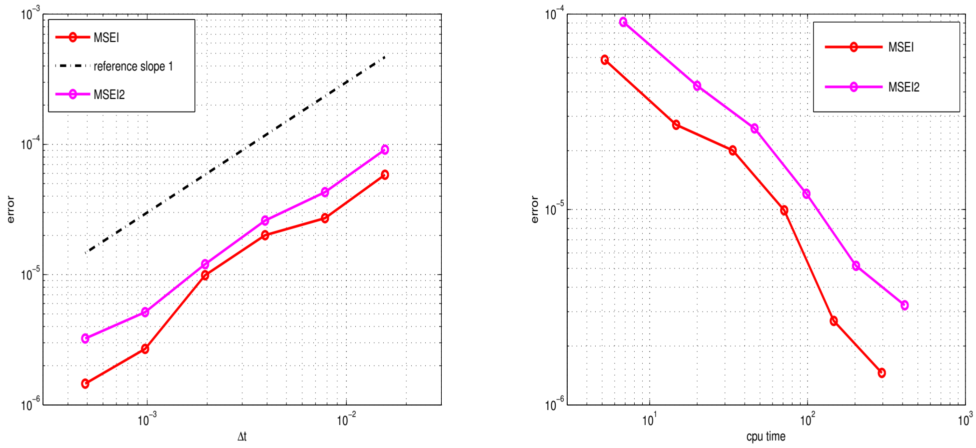


Figure 2.12: Left figure shows strong convergence for the MSEI and MSEI2 schemes, both approximating solution to 3-D Langevin system with commutative noise. The slope of the best-fit line for MSEI is 1.075582 while for MSEI2 is 0.982340. Right figure shows the numerical error against computer process time.

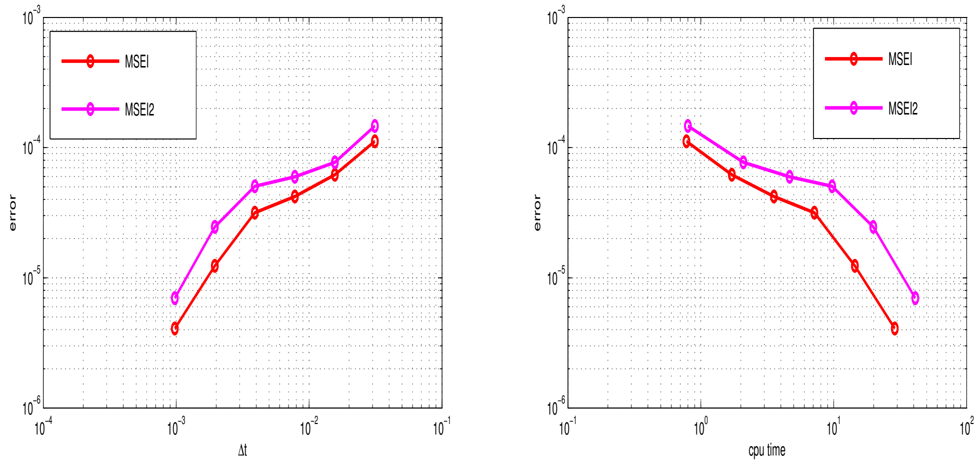


Figure 2.13: Strong convergence for MSEI and MSEI2 schemes on the left, both when approximating solution to 3-D Langevin system with general i.e. non-diagonal, non-commutative noise. Slope of the best-fit line for MSEI is 0.839509 while for MSEI2 is 0.775424. Right figure shows the numerical error against computer process time.

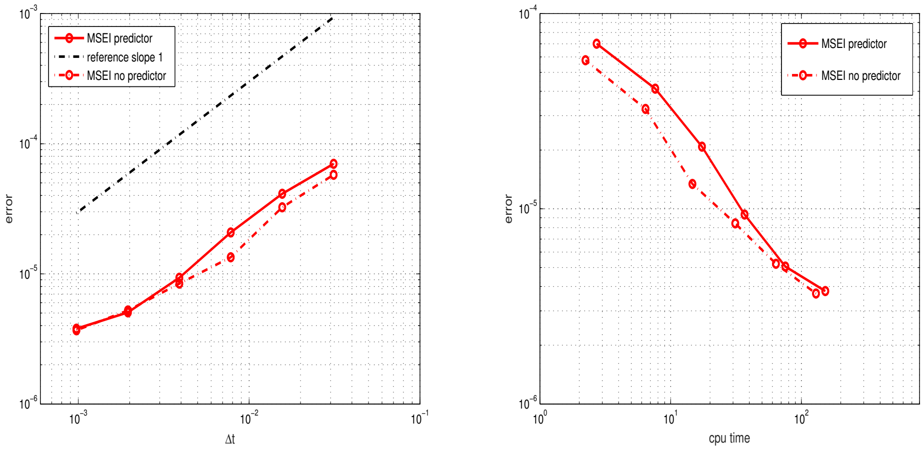


Figure 2.14: Comparison between the cases where we use a predictor and where we do not use one. The slope of the fitted line for the MSEI predictor case is 0.893925, while for the MSEI no-predictor is 0.812193. 3-d LLG system with non-diagonal, non-commutative noise. Reference time step-size is 2^{-12} and we have considered 500 samples.

In the analysis that follows, instead of approximating $\tilde{\mathbf{u}}_{n+1}$ by a numerical scheme, we obtain it by Taylor-expanding $G(\tilde{\mathbf{u}}(t_{n+1}))$ about $\mathbf{u}(t_n)$. Then, (2.26) becomes

$$\mathbf{u}(t_{n+1}) = e^{\Delta t A}\mathbf{u}(t_n) + \mathbf{F}(\mathbf{u}(t_n))\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}ds + \int_{t_n}^{t_{n+1}} e^{\frac{\Delta t A}{2}} G(\mathbf{u}(t_n))d\mathbf{W}(s) \tag{2.28}$$

$$+ \frac{1}{2}\int_{t_n}^{t_{n+1}} e^{\frac{\Delta t A}{2}} DG(\mathbf{u}(t_n))\big(\mathbf{u}(t_{n+1}) - \mathbf{u}(t_n)\big)d\mathbf{W}(s)$$

$$+ \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A} DG(\mathbf{u}(t_n))\Big(G(\mathbf{u}(t_n))\int_{t_n}^{s} d\mathbf{W}(\tau)\Big) \circ d\mathbf{W}(s)$$

$$+ R_T(s; t_n, \mathbf{u}(t_n)) + R_M(s; t_n, \mathbf{u}(t_n)),$$

where by R_T we denote the remainder higher order term from the Taylor expansion. Note that R_T has the same order with the the second term of R_M as given in equation (2.25).

For the proof of convergence, following the same pattern with Subsection 2.1, we use the following version of the approximate solution which is just a recurrence of (2.28).

$$\mathbf{u}_n = e^{t_n A}\mathbf{u}_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{F}(\mathbf{u}_k)ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}G(\mathbf{u}_k)d\mathbf{W}(s) \tag{2.29}$$

$$+\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}DG(\mathbf{u}_k)\big(\mathbf{u}_{k+1}-\mathbf{u}_k\big)d\mathbf{W}(s)$$

$$+\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}DG(\mathbf{u}_k)\Big(G(\mathbf{u}_k)\int_{t_k}^{s} d\mathbf{W}(\tau)\Big)\circ d\mathbf{W}(s)$$

$$+\sum_{k=0}^{n-1} R_T(s;t_k,\mathbf{u}_k) + \sum_{k=0}^{n-1} R_M(s;t_k,\mathbf{u}_k).$$

Now, recall that the exact solution, after we apply the Itô-Stratonovich correction formula satisfies

$$\mathbf{u}(t_n) = e^{t_n A}\mathbf{u}_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{F}(\mathbf{u}(s))ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}G(\mathbf{u}(s))d\mathbf{W}(s)$$

$$+\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s))ds,$$

where $\mathbf{c}(\mathbf{u})$ is a vector in $\mathbb{R}^d$ as defined in Assumption 2.1.2. After we Taylor-expand $G(\mathbf{u}(s))$ about $\mathbf{u}_k$ in the subintervals $[t_k, t_{k+1}]$, the exact solution of (1.38) is given by

$$\begin{aligned}\mathbf{u}(t_n) = e^{t_n A}\mathbf{u}_0 &+ \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{F}(\mathbf{u}(s))ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}G(\mathbf{u}_k)d\mathbf{W}(s)\\ &+\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}DG(\mathbf{u}_k)\big(\mathbf{u}(s)-\mathbf{u}_k\big)d\mathbf{W}(s)\\ &+\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s))ds + \sum_{k=0}^{n-1} R_T(s;t_k,\mathbf{u}_k).\end{aligned} \tag{2.30}$$

Finally, by substituting the difference $\mathbf{u}(s)-\mathbf{u}_k$, the exact solution at time t_n is given by

$$
\mathbf{u}(t_n) = e^{t_n A}\mathbf{u}_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{F}(\mathbf{u}(s))ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}G(\mathbf{u}_k)d\mathbf{W}(s) \tag{2.31}
$$

$$
+ \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}DG(\mathbf{u}_k)\Big(G(\mathbf{u}_k)\int_{t_k}^{s} d\mathbf{W}(\tau)\Big)d\mathbf{W}(s)
$$

$$
+ \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}DG(\mathbf{u}_k)\big(\mathbf{F}(\mathbf{u}_k)(t_k-s)+R_E\big)d\mathbf{W}(s)
$$

$$
+ \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s))ds + \sum_{k=0}^{n-1} R_T(s;t_k,\mathbf{u}_k).
$$

As for all the previous schemes, we aim at considering the difference between the exact and the approximate solution, i.e. (2.31)-(2.29) and grouping it in terms that are convenient for bounding separately. In that direction, we firstly 'pair' the drift and diffusion terms and then we notice that the double integral in (2.29) is a Stratonovich integral while the double integral in (2.31) is an Itô one. If we 'correct' the Stratonovich integral, expand them both and subtract them, after the cancellations that occur, we are left with $-\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{c}(\mathbf{u}_k)ds$. Next, we group this with the sixth term of (2.31) to get

$$
I := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\big(\mathbf{F}(\mathbf{u}(s))-\mathbf{F}(\mathbf{u}_k)\big)ds
$$

$$
II := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \big(e^{(t_n-s)A}-e^{(t_n-t_{k+\frac{1}{2}})A}\big)G(\mathbf{u}_k)d\mathbf{W}(s)
$$

$$
III := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\big(\mathbf{c}(\mathbf{u}(s))-\mathbf{c}(\mathbf{u}_k)\big)ds
$$

$$
IV := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}DG(\mathbf{u}_k)\big(\mathbf{u}_{k+1}-\mathbf{u}_k\big)d\mathbf{W}(s)
$$

$$
V := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} (e^{(t_n-t_{k+\frac{1}{2}})A}-e^{(t_n-s)A})\int_0^1(1-\theta)D^2G(\mathbf{u}_k+\theta(\mathbf{u}(s)-\mathbf{u}_k))\big(\mathbf{u}(s)-\mathbf{u}_k\big)^2 d\theta d\mathbf{W}(s)
$$

$$+\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}(e^{(t_n-t_{k+\frac{1}{2}})A}-e^{(t_n-s)A})DG(\mathbf{u}_k)\big(\mathbf{F}(\mathbf{u}_k)(t_k-s)+R_E\big)d\mathbf{W}(s)$$

$$+\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}e^{(t_n-t_{k+\frac{1}{2}})A}\int_0^1(1-\theta)D^2G(\mathbf{u}_k+\theta(\mathbf{u}_{k+1}-\mathbf{u}_k))(\mathbf{u}_{k+1}-\mathbf{u}_k)^2d\theta d\mathbf{W}(s)$$

$$+\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}e^{(t_n-t_{k+\frac{1}{2}})A}D\mathbf{F}(\mathbf{u}_k)\big(\mathbf{u}(s)-\mathbf{u}_k\big)ds.$$

The upper bound for the first term is given by the inequality

$$\mathbb{E}[\|I\|_2^2]\leqslant C_1\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}[\|\mathbf{u}(s)-\mathbf{u}_k\|_2^2]ds,$$

with C_1 being a positive constant independent of Δt. This result and its proof are given in more detail in Lemma 2.1.3.

Lemma 2.3.1. *Let Assumptions 1.1.9 and 1.1.10 hold. Then*

$$\mathbb{E}\big[\|II\|_2^2\big]\leqslant C_2\Delta t^{1+2\gamma},$$

where $\gamma\in(\frac{1}{2},1)$ and C_2 is a positive constant independent of Δt.

Proof. By Itô's isometry, Remark 2.1.4 and Proposition 1.4.9 we get

$$\mathbb{E}[\|II\|_2^2]=\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\big[\|\big(e^{(t_n-s)A}-e^{(t_n-t_{k+\frac{1}{2}})A}\big)G(\mathbf{u}_k)\|_F^2\big]ds$$

$$\leqslant\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\|(-A)^{-\gamma}\big(e^{(t_n-s)A}-e^{(t_n-t_{k+\frac{1}{2}})A}\big)\|_F^2\mathbb{E}[\|(-A)^{\gamma}G(\mathbf{u}_k)\|_F^2]ds$$

$$\leqslant C\int_{t_{n-1}}^{t_n}\|(-A)^{-\gamma}\big(e^{(t_n-s)A}-e^{(\frac{\Delta tA}{2})}\big)\|_F^2ds$$

$$+\sum_{k=0}^{n-2}\int_{t_k}^{t_{k+1}}\|(-A)^{-\gamma}\big(e^{(s-t_{k+\frac{1}{2}})A}-I\big)\|_F^2\|(-A)^{1-\gamma}e^{(t_n-s)A}\|_F^2ds$$

$$\leqslant C\int_{t_{n-1}}^{t_n}\|(-A)^{-\gamma}\big(e^{(t_n-s)A}-e^{\frac{\Delta tA}{2}}\big)\|_F^2ds$$

$$+\Delta t^2\sum_{k=0}^{n-2}\int_{t_k}^{t_{k+1}}\|(-A)^{1-\gamma}e^{(t_n-s)A}\|_F^2ds$$

$$\leqslant C\int_{t_{n-1}}^{t_n}(s-t_{n-\frac{1}{2}})^{2\gamma}ds+\Delta t^2\sum_{k=0}^{n-2}\int_{t_k}^{t_{k+1}}(t_n-s)^{2(\gamma-1)}ds$$

$$\leqslant C\Delta t^{(1+2\gamma)}+\Delta t^3\Big(\sum_{k=0}^{n-2}\int_{t_k}^{t_{k+1}}(t_n-t_{k+\frac{1}{2}})^{2(\gamma-1)}\Delta t^{2(\gamma-1)}\Big)$$

$$\leqslant C\Delta t^{(1+2\gamma)}+\Delta t^{(1+2\gamma)}\sum_{k=0}^{n-1}t_k^{2(\gamma-1)}\leqslant K\Delta t^{(1+2\gamma)}\big(2+\frac{1}{1-2\gamma}\big)$$

$$\leqslant C_2\Delta t^{(1+2\gamma)}.$$

□

The above term II is similar with the term II_2 from Section 2.1. The proof followed here, as seen in [37], is essentially an alternative to the proof of Lemma 2.1.5.

Lemma 2.3.2. *Let Assumptions 1.1.9, 1.1.10 and 2.1.2 hold. Then*

$$\mathbb{E}\big[\|III\|_2^2\big]\leqslant C_3\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}[\|\boldsymbol{u}(s)-\boldsymbol{u}_k\|_2^2]$$

where C_3 is a positive constant independent of Δt.

Proof. An upper bound for this term is found in a similar way with the term $\mathbb{E}[\|I\|_2^2]$. For more details see proof of Lemmas 2.1.3 and 2.1.6. □

Lemma 2.3.3. *Let Assumptions 1.1.9-1.4.8 and 2.1.2 hold. Then*

$$\mathbb{E}\big[\|IV\|_2^2\big]\leqslant C_4\Delta t^2,$$

where C_4 is a positive constant independent of Δt.

Proof. For this term we use Itô's isometry, Assumption 2.1.2, the boundedness of the exponential operator and we take $\mathbf{u}_{k+1}$ by equation (1.23). Then,

$$\mathbb{E}[\|IV\|_2^2]=\frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\big[\|e^{(t_n-t_{k+\frac{1}{2}})A}DG(\mathbf{u}_k)\big(\mathbf{u}_{k+1}-\mathbf{u}_k\big)\|_F^2\big]ds$$

$$\leqslant C\sup_{0\leqslant k\leqslant n-1}\mathbb{E}\big[(1+\|\mathbf{u}_k\|_2^2)\big]\Bigg(\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}[\|e^{(t_n-t_{k+\frac{1}{2}})A}(e^{\Delta tA}-I)\mathbf{u}_k\|_2^2]ds$$

$$+\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\big[\|e^{(t_n-t_{k+\frac{1}{2}})A}\int_{t_k}^{s}e^{(s-\tau)A}\mathbf{F}(\mathbf{u}(\tau))d\tau\|_2^2\big]ds$$

$$+\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}[\|e^{(t_n-t_{k+\frac{1}{2}})A}\int_{t_k}^{s}e^{(s-\tau)A}G(\mathbf{u}(\tau))d\mathbf{W}(\tau)\|_2^2]ds\Bigg).$$

We seek upper bounds for the last three terms above separately. For the first term by Proposition 1.4.9 we have

$$\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}[\|(-A)^{\gamma}e^{(t_n-t_{k+\frac{1}{2}})A}(-A)^{-\gamma}(e^{\Delta tA}-I)\mathbf{u}_k\|_2^2]ds$$

$$\leqslant \sup_{0\leqslant k\leqslant n-1}\mathbb{E}[\|\mathbf{u}_k\|_2^2]\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\Big(\frac{\Delta t}{t_n-t_{k+\frac{1}{2}}}\Big)^{2\gamma}ds$$
$$\leqslant C_{4,1}\Delta t^{2\gamma+1}.$$

where $C_{4,1}$ is a positive constant independent of Δt. Similarly for the second term,

$$\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\big[\|e^{(t_n-t_{k+\frac{1}{2}})A}\int_{t_k}^{s}e^{(s-\tau)A}\mathbf{F}(\mathbf{u}(\tau))d\tau\|_2^2\big]ds$$

$$\leqslant \sup_{0\leqslant k\leqslant n-1}\mathbb{E}[\|e^{(t_n-t_{k+\frac{1}{2}})A}\|_F^2]\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\big[\|\int_{t_k}^{s}e^{(s-\tau)A}\mathbf{F}(\mathbf{u}(\tau))d\tau\|_2^2\big]ds$$

$$\leqslant \sup_{0\leqslant k\leqslant n-1}\mathbb{E}[\|e^{(t_n-t_{k+\frac{1}{2}})A}\|_F^2]\sup_{0\leqslant s\leqslant t_n}\sup_{t_k\leqslant\tau\leqslant s}\mathbb{E}[\|e^{(s-\tau)A}\mathbf{F}(\mathbf{u}(\tau))\|_2^2]\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}(s-t_k)^2ds$$

$$\leqslant \sup_{0\leqslant k\leqslant n-1}\mathbb{E}[\|e^{(t_n-t_{k+\frac{1}{2}})A}\|_F^2]\sup_{0\leqslant s\leqslant t_n}\sup_{t_k\leqslant\tau\leqslant s}\mathbb{E}[\|e^{(s-\tau)A}\mathbf{F}(\mathbf{u}(\tau))\|_2^2]\Delta t^3=C_{4,2}\Delta t^3.$$

Finally, for the third term

$$\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}[\|e^{(t_n-t_{k+\frac{1}{2}})A}\int_{t_k}^{s}e^{(s-\tau)A}G(\mathbf{u}(\tau))d\mathbf{W}(\tau)\|_2^2]ds$$

$$\leqslant \sup_{0\leqslant k\leqslant n-1}\mathbb{E}[\|e^{(t_n-t_{k+\frac{1}{2}})A}\|_F^2]\sup_{0\leqslant s\leqslant t_n}\sup_{t_k\leqslant\tau\leqslant s}\mathbb{E}[\|e^{(s-\tau)A}\|_F^2]$$
$$\times\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\big[\|\int_{t_k}^{s}G(\mathbf{u}(\tau))d\mathbf{W}(\tau)\|_2^2\big]ds$$

$$\leqslant \sup_{0\leqslant k\leqslant n-1}\mathbb{E}[\|e^{(t_n-t_{k+\frac{1}{2}})A}\|_F^2]\sup_{0\leqslant s\leqslant t_n}\sup_{t_k\leqslant\tau\leqslant s}\mathbb{E}[\|e^{(s-\tau)A}\|_F^2]\Delta t^2=C_{4,3}\Delta t^2.$$

Using these individual estimates, we get

$$\mathbb{E}[\|IV\|_2^2 \leqslant C_{4,1}\Delta t^{2\gamma+1} + C_{4,2}\Delta t^3 + C_{4,3}\Delta t^2.$$

The observation $\Delta t^2 \geqslant \Delta t^{1+2\gamma}$, for $\gamma > \frac{1}{2}$ gives the final estimate. □

Theorem 2.3.4.
Under the Assumptions 1.1.9, 1.1.10, 1.4.8 and 2.1.2, let $\boldsymbol{u}_n$ as represented by equation (2.29) be the approximate solution of the semilinear Stratonovich SDE (1.38) at time $t_n = n\Delta t$, $\Delta t > 0$. Also, let $\boldsymbol{u}(t_n)$, as given in (2.31), denote the exact solution of the same Stratonovich-interpreted SDE. Then

$$\left(\mathbb{E}\left[\|\boldsymbol{u}(t_n) - \boldsymbol{u}_n\|_2^2\right]\right)^{1/2} \leqslant C\Delta t,$$

where C is a positive constant, independent of Δt.

Proof. Putting the individual estimates from Lemmas 2.1.3 and 2.3.1-2.3.3 together, we get

$$\mathbb{E}\left[\|\mathbf{u}(t_n) - \mathbf{u}_n\|_2^2\right] \leqslant K\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}[\|\mathbf{u}(s) - \mathbf{u}_k\|_2^2]ds + L\Delta t^{1+2\gamma} + C_4\Delta t^2,$$

where $K = \max\{C_1, C_{3,1}\}$, $L = \max\{C_2, C_{3,2}\}$ and C_4 as in Lemma 2.3.3. Recall that $\Delta t^2 \geqslant \Delta t^{1+2\gamma}$, for $\gamma > \frac{1}{2}$, hence

$$\mathbb{E}\left[\|\mathbf{u}(t_n) - \mathbf{u}_n\|_2^2\right] \leqslant K\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}[\|\mathbf{u}(s) - \mathbf{u}_k\|_2^2]ds + M\Delta t^2,$$

where $M = \max\{L, C_4\}$. Then, by Gronwall's lemma we get

$$\mathbb{E}\left[\|\mathbf{u}(t_n) - \mathbf{u}_n\|_2^2\right] \leqslant K\Delta t^2\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{K\Delta t(t-s)}ds + M\Delta t^2.$$

Since $\displaystyle\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{K\Delta t(t-s)}ds$ is bounded we have that

$$\mathbb{E}\left[\|\mathbf{u}(t_n) - \mathbf{u}_n\|_2^2\right] \leqslant C^2\Delta t^2.$$

Taking square roots on both sides finishes the proof. □

In this Chapter, we proved analytically and examined numerically strong convergence for the SEI and the new MSEI scheme, both suited for Stratonovich SDEs. In Chapter 3, we derive a new version of SEI that enables a generalisation of its application to SDEs which are interpreted by any stochastic calculus. Then, we prove that the generalised version of SEI converges strongly and that the rate of convergence depends on the choice of the supporting value $\tilde{\mathbf{u}}_{n+1}$.

Chapter 3

Generalised version of SEI scheme

In Subsection 1.2.3 and Section 2.1 we analysed an exponential based scheme for SDEs interpreted in the Itô and the Stratonovich sense. It turns out though that in some cases neither the Itô nor the Stratonovich integral is a suitable interpretation, see for example [17, 18, 48, 65]. Under this observation, in this chapter we consider a generalisation of the stochastic integral interpretation. Then, we extend the exponential scheme of Section 2.1, so that it applies to the new general context. Moreover, we provide a general expression for the predictor of the SEI scheme. This enables us to experiment with obtaining the predictor from explicit schemes of different orders and see how these orders affect the overall order of convergence for SEI.

3.1 Generalisation of the stochastic integral interpretation

Let us consider the one-dimensional general SDE of the form

$$du = F(u)dt + G(u) * dW, \tag{3.1}$$

which is equivalently given by the following integral equation in a time interval $[0, t]$

$$u(t) = u(0) + \int_0^t F(u(s))ds + \int_0^t G(u(s)) * dW(s). \tag{3.2}$$

In this chapter, we are interested in defining the generalised version of the stochastic integral on the right hand side of equation (3.2). Moreover, we derive a scheme for approximating solutions to equations of the form (3.2). Taking into consideration the fact that the definition of the Itô and the Stratonovich integrals use the left-end point and the mid-point respectively for evaluating the integrable function, we examine the generalisation of the stochastic integral interpretation, starting from the following one-dimensional definition which is extended to higher dimensions below.

Definition 3.1.1. ***One-dimensional general stochastic integral*** *[77]*

Let u be a real-valued stochastic process in $L^2(\Omega, \mathbb{R})$ and let $\phi : \mathbb{R} \times [0,T] \to \mathbb{R}$ $\in C^2(\mathbb{R})$ such that $\mathbb{E}\Big(\int_0^T |\phi(u,s)|^2 ds\Big) < \infty$. Then,

$$\int_0^T \phi(u(s), s) * dW(s) := \lim_{N\to\infty} \sum_{j=0}^{N-1} \phi(\tilde{z}_j, t_j)\Delta W_j,$$

where $t_j := j\Delta t$, $\tilde{z}_j = (1-\alpha)u(t_j) + \alpha u(t_{j+1})$, $\Delta W_j = W(t_{j+1}) - W(t_j)$ are Brownian increments and $\alpha \in [0,1]$.

Note that for $\alpha = \frac{1}{2}$ the above definition gives $\tilde{z}_j = \frac{1}{2}\big(u(t_j) + u(t_{j+1})\big)$, which corresponds to the point choice for the Stratonovich integral, while $\alpha = 0$ gives $\tilde{z}_j = u(t_j)$, that is the left-end point which corresponds to the Itô integral. We are also interested in finding a connection between the different interpretations, similar to the Itô-Stratonovich correction formula. Recall from equations (1.4) and (1.7) of Chapter 1 for example, that for $d = m = 1$, if u is a solution of the Itô SDE $du = f(u)dt + G(u)dW$, then

$$\int_0^t \phi(u(s)) \circ dW(s) = \int_0^t \phi(u(s))dW(s) + \frac{1}{2}[\phi(u), W] \tag{3.3}$$

where $[\phi(u), W]$ is the quadratic variation of ϕ defined by

$$[\phi(u), W] := \lim_{N\to\infty} \sum_{j=0}^{N-1} \big(\phi(u(t_{j+1})) - \phi(u(t_j))\big)\Delta W_j = \int_0^t \phi'(u(s))G(u(s))ds.$$

A generalisation of this correction formula is given in the following proposition.

Proposition 3.1.2. *Let $\phi \in C^2(\mathbb{R})$, u be a solution of (3.1) and let $\alpha \in [0,1]$. Then,*

$$\int_0^t \phi(u(s)) * dW(s) = \int_0^t \phi(u(s))dW(s) + \alpha[\phi(u), *W]$$

where $[\phi(u), *W]$ is the generalised version of the quadratic variation, defined as

$$[\phi(u), *W] := \lim_{N\to\infty} \sum_{j=0}^{N-1} \big((1-\alpha)\phi(u(t_j)) - \alpha\phi(u(t_{j+1}))\big)\Delta W_j = \int_0^t \phi'(u(s))G(u(s))ds.$$

For more details about the generalised version of the correction formula, we refer to [77], where a discrete version of the formula has been derived and proved. The corresponding continuous version can be easily obtained from the discrete, just by taking the limit as the mesh size goes to zero.

A general discussion on correction formulas for stochastic integrals other than the Itô and Stratonovich, for example for the Marcus stochastic integral, can be found in [3].

An alternative proof for Proposition 3.1.2 can be based on the Mean Value Theorem and the Taylor expansion and is similar to the one for the Itô-Stratonovich correction formula. A sketch of it for $d = m = 1$ follows the basic steps of the latter as seen in [42] and can be found in Section 7.2 of the Appendix.

In a similar way with the one-dimensional case, we give the definition of the general stochastic integral in higher dimensions.

Definition 3.1.3. ***General stochastic integral in higher dimensions***
Let $\boldsymbol{u}$ be a real-valued stochastic process in $L^2(\Omega, \mathbb{R}^d)$ and let $\phi : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d\times m}$ be a C^2 function such that $\mathbb{E}\big(\int_0^T |\phi(\boldsymbol{u}, s)|^2 ds\big) < \infty$. Then,

$$\int_0^T \phi(\boldsymbol{u}(s), s) * d\,\boldsymbol{W}(s) = \int_0^T \begin{pmatrix} \phi_{11}(\boldsymbol{u}, s) & \dots & \phi_{1m}(\boldsymbol{u}, s) \\ \vdots & \ddots & \vdots \\ \phi_{d1}(\boldsymbol{u}, s) & \dots & \phi_{dm}(\boldsymbol{u}, s) \end{pmatrix} * \begin{pmatrix} dW_1 \\ \vdots \\ dW_m \end{pmatrix}$$

is a $d \times 1$ column vector with i-th component given by

$$\sum_{\ell=1}^{m} \int_0^T \phi_{i\ell}(\boldsymbol{u}, s) * dW_\ell(t) := \sum_{\ell=1}^{m} \lim_{n\to\infty} \sum_{j=0}^{n-1} \phi_{i,\ell}(\tilde{\boldsymbol{z}}, t_j)\Delta W_{\ell,j}, \;\; i = 1, \dots, d \tag{3.4}$$

where $\Delta W_{\ell,j} = W_\ell(t_{j+1}) - W_\ell(t_j)$, $\ell = 1, \dots, m$ are Brownian increments, $t_j := j\Delta t$, $d, m \in \mathbb{Z}$ and $\tilde{\boldsymbol{z}}_j = (1-\alpha)\boldsymbol{u}(t_j) + \alpha\boldsymbol{u}(t_{j+1})$, $\alpha \in [0, 1]$.

Note that the terms of the sum in the RHS of equation (3.4) are components of the term $\phi(\rho)\Delta W_j$, $j = 0, \dots, N-1$ of equation (7.5) in the sketch of the proof for Proposition 3.1.2.

Now, let us consider the higher-dimensional general SDE

$$d\mathbf{u} = \big[A\mathbf{u} + \mathbf{F}(\mathbf{u})\big]dt + G(\mathbf{u}) * d\mathbf{W}, \tag{3.5}$$

or in integral form in the time interval $[0, t]$

$$\mathbf{u}(t) = \int_0^t \big[A\mathbf{u}(s) + \mathbf{F}(\mathbf{u}(s))\big]ds + \int_0^t G(\mathbf{u}(s)) * d\mathbf{W}(s), \tag{3.6}$$

where $A \in \mathbb{R}^{d\times d}$ satisfies Assumption 1.4.8 and $\mathbf{F} : \mathbb{R}^d \to \mathbb{R}^d$, $G : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are Lipschitz functions with linear growth, that is they satisfy Assumptions 1.1.9 and

1.1.10. Also $\mathbf{W}$ is an m-dimensional Wiener process of the form $(W_1, \dots, W_m)^\top$. The stochastic integral of equation (3.6) is evaluated according to [77], using the following analogue of Proposition 3.1.2 for higher dimensions. Let $\phi : \mathbb{R}^d \times [0, T] \to \mathbb{R}^{d\times m} \in C^2(\mathbb{R}^d)$, then for a solution $\mathbf{u}$ of (3.5) we have

$$\int_0^t \phi(\mathbf{u}(s)) * d\mathbf{W}(s) = \int_0^t \phi(\mathbf{u}(s))d\mathbf{W}(s) + \alpha \int_0^t \mathbf{c}(\mathbf{u}(s))ds, \tag{3.7}$$

where $\mathbf{c}(\mathbf{u}) = (c_1(\mathbf{u}), \dots, c_d(\mathbf{u}))^\top$ is a vector in $\mathbb{R}^d$, the i-th component of which is given by

$$c_i = \sum_{k=1}^{d}\sum_{j=1}^{d} G_{kj}\frac{\partial \phi_{ij}(\mathbf{u})}{\partial u_k}, \quad i = 1, \dots, d. \tag{3.8}$$

We consider the variational formulation of equation (3.5). Firstly, we multiply both sides of (3.5) by e^{-tA}

$$e^{-tA}d\mathbf{u} = e^{-tA}\big[A\mathbf{u} + \mathbf{F}(\mathbf{u})\big]dt + e^{-tA}G(\mathbf{u}) * d\mathbf{W}.$$

Then, by rearranging and integrating over the interval $[t_n, t_{n+1}]$, we get that the solution $\mathbf{u}(t)$ of (3.5) satisfies

$$\mathbf{u}(t_{n+1}) = e^{\Delta t A}\mathbf{u}(t_n) + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{F}(\mathbf{u}(s))ds + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}G(\mathbf{u}(s)) * d\mathbf{W}(s). \tag{3.9}$$

On the right hand side of (3.9) we have a deterministic integral and a general stochastic integral that can be interpreted in a way suggested by the following definition.

3.2 Generalisation of SEI and strong convergence analysis

Similar to the Itô and the Stratonovich cases, in order to derive the exponential integrators-based scheme, we start from evaluating $F(\mathbf{u}(s))$ at the left-end point of the intervals $[t_n, t_{n+1}]$ and then we compute the deterministic integral of (3.9) in the following way

$$\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{F}(\mathbf{u}(t_n))ds = A^{-1}(e^{\Delta t A} - I)\mathbf{F}(\mathbf{u}(t_n)).$$

For the general stochastic integral, we use the following approximation

$$\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}G(\mathbf{u}(s)) * d\mathbf{W}(s) \approx$$

$$\big[(1-\alpha)e^{(1-\alpha)\Delta tA}G(\mathbf{u}(t_n)) + \alpha e^{\alpha\Delta tA}G(\mathbf{u}(t_{n+1}))\big]\Delta\mathbf{W}_n,$$

for $\alpha \in [0,1]$. This approximation, α in particular, reflects the choice of the intermediate point at which the integrand is evaluated. See for example that for $\alpha = 0$ we get $e^{\Delta tA}G(\mathbf{u}(t_n))\Delta\mathbf{W}_n$, which corresponds to the approximation of the Itô stochastic integral using the left-end point of the interval $[t_n, t_{n+1}]$. For $\alpha = \frac{1}{2}$ on the other hand, we get $\frac{1}{2}e^{\frac{\Delta tA}{2}}\big(G(\mathbf{u}(t_n)) + G(\mathbf{u}(t_{n+1}))\big)\Delta\mathbf{W}_n$, which corresponds to the approximation of the Stratonovich integral using the mid-point of each subinterval, as in (1.42) of Section 1.2.6. Then, our scheme becomes

$$\mathbf{u}_{n+1} = e^{\Delta tA}\mathbf{u}_n + A^{-1}(e^{\Delta tA} - I)\mathbf{F}(\mathbf{u}_n) + \big[(1-\alpha)e^{(1-\alpha)\Delta tA}G(\mathbf{u}_n) + \alpha e^{\alpha\Delta tA}G(\tilde{\mathbf{u}}_{n+1})\big]\Delta\mathbf{W}_n \tag{3.10}$$

where $\mathbf{u}_{n+1}$ denotes the approximate solution at time t_{n+1} and $\tilde{\mathbf{u}}_{n+1}$ is a predictor given, in general, by the expression

$$\tilde{\mathbf{u}}_{n+1} = S_1\mathbf{u}_n + S_2\mathbf{F}(\mathbf{u}_n) + S_3G(\mathbf{u}_n)\Delta\mathbf{W}_n. \tag{3.11}$$

Throughout the analysis and experiments of the previous chapters, we have used different methods like the Euler-Maruyama method as well as the exponential integrators for Itô (ETD) in order to obtain $\tilde{\mathbf{u}}_{n+1}$. The idea behind this general version of the predictor is that we can obtain all the previously mentioned or any other method by substituting in (3.11) for S_1, S_2, S_3 accordingly. The connection between the predictors obtained by different methods, is related to the order of convergence of each method. In the following examples, we see predictors taken by the exponential integrators and the implicit Euler-Maruyama schemes.

Example 1. *Let $\tilde{\boldsymbol{u}}_{n+1}$ be given by the exponential integrators scheme as given in (1.24). Then $S_1 = S_3 = e^{\Delta tA}$ and $S_2 = A^{-1}(e^{\Delta tA} - I)$, so that*

$$\tilde{\boldsymbol{u}}_{n+1} = e^{\Delta tA}\boldsymbol{u}_n + A^{-1}(e^{\Delta tA} - I)\boldsymbol{F}(\boldsymbol{u}_n) + e^{\Delta tA}G(\boldsymbol{u}_n)\Delta\boldsymbol{W}_n.$$

By Taylor expansion of the exponential series we have

$$e^{\Delta tA} = I + \Delta tA + \frac{1}{2}\Delta t^2A^2 + \dots$$
$$A^{-1}(e^{\Delta tA} - I) = \Delta t + \frac{1}{2}\Delta t^2A + \dots$$

Hence,

$$\tilde{\boldsymbol{u}}_{n+1} = (I + \Delta tA)\boldsymbol{u}_n + \Delta t\boldsymbol{F}(\boldsymbol{u}_n) + (I + \Delta tA)G(\boldsymbol{u}_n)\Delta\boldsymbol{W}_n + h.o.t.$$

Example 2. *In this example we consider the predictor $\tilde{\boldsymbol{u}}_{n+1}$ to be taken by the* *implicit Euler-Maruyama scheme, where* $S_1 = S_3 = (I - \Delta tA)^{-1}$ *and* $S_2 = (I - \Delta tA)^{-1}\Delta t$. *Then,*

$$\tilde{\boldsymbol{u}}_{n+1} = (I - \Delta tA)^{-1}\boldsymbol{u}_n + (I - \Delta tA)^{-1}\boldsymbol{F}(\boldsymbol{u}_n)\Delta t + (I - \Delta tA)^{-1}G(\boldsymbol{u}_n)\Delta \boldsymbol{W}_n.$$

Again by Taylor expansion

$$(I - \Delta tA)^{-1} = I + \Delta tA + \dots$$

$$\Delta t(I - \Delta tA) = \Delta t + \Delta t^2 A + \dots$$

Hence,

$$\tilde{\boldsymbol{u}}_{n+1} = (I + \Delta tA)\boldsymbol{u}_n + \Delta t\boldsymbol{F}(\boldsymbol{u}_n) + (I + \Delta tA)G(\boldsymbol{u}_n)\Delta \boldsymbol{W}_n + h.o.t$$

We see that the approximations in Examples 1 and 2 are of the same order. We conclude that as long as the method used to obtain the predictor $\tilde{\mathbf{u}}_{n+1}$ is of a specific known order of convergence, we can impose the following assumptions on S_1, S_2, S_3.

Assumption 3.2.1. *Let* $S = e^{\Delta tA}$, *then let* S_1, S_2 *and* S_3, *as in (3.11), satisfy the following inequalities*

$$\|S - S_\ell\|_F \leqslant D\Delta t^p, \quad \ell = 1, 3$$

and

$$\|S_2\|_F \leqslant D\Delta t^q,$$

where $D, p, q > 0$ *are constants.*

In preparation for the strong convergence analysis, we now write the scheme of equation (3.10) in a continuous form. In order to derive this form, we substitute for the predictor $\tilde{\mathbf{u}}_{n+1}$ and Taylor expand $G(\tilde{\mathbf{u}}_{n+1})$ about $\mathbf{u}_n$ in the interval $[t_n, t_{n+1}]$. So, since $S_1\mathbf{u}_n = \mathbf{u}_n + (S_1 - I)\mathbf{u}_n$,

$$\begin{aligned} G(\tilde{\mathbf{u}}_{n+1}) &= G(\mathbf{u}_n) + DG(\mathbf{u}_n)\big[(S_1 - I)\mathbf{u}_n + S_2\mathbf{F}(\mathbf{u}_n) + S_3G(\mathbf{u}_n)\Delta\mathbf{W}_n\big] + R \\ &= G(\mathbf{u}_n) + DG(\mathbf{u}_n)\hat{\mathbf{h}} + R \end{aligned}$$

where $\hat{\mathbf{h}} := (S_1 - I)\mathbf{u}_n + S_2\mathbf{F}(\mathbf{u}_n) + S_3G(\mathbf{u}_n)\Delta\mathbf{W}_n$ and R is the second order remainder term of the expansion given in integral form by

$$R := \frac{1}{2}\int_0^1 (1 - \theta)\hat{\mathbf{h}}^\top D^2G(\mathbf{u}_n + \theta\hat{\mathbf{h}})\hat{\mathbf{h}}d\theta. \tag{3.12}$$

Thus, the continuous version of the numerical approximation to equation (3.5) is given by

$$
\begin{aligned}
\mathbf{u}_{n+1} = e^{\Delta t A}\mathbf{u}_n &+ \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{F}(\mathbf{u}_n)ds \\
&+ \int_{t_n}^{t_{n+1}} \big((1-\alpha)e^{(1-\alpha)\Delta t A}G(\mathbf{u}_n) + \alpha e^{\alpha\Delta t A}G(\mathbf{u}_n)\big)d\mathbf{W}(s) \\
&+ \int_{t_n}^{t_{n+1}} \alpha e^{\Delta t A}DG(\mathbf{u}_n)(S_1 - I)\mathbf{u}_n dW(s) + \int_{t_n}^{t_{n+1}} \alpha e^{\alpha\Delta t A}DG(\mathbf{u}_n)S_2\mathbf{F}(\mathbf{u}_n)\Delta t d\mathbf{W}(s) \\
&+ \int_{t_n}^{t_{n+1}} \alpha e^{\alpha\Delta t A}S_3\mathbf{c}(\mathbf{u}_n)ds + \frac{1}{2}\int_{t_n}^{t_{n+1}} \alpha e^{\alpha\Delta t A}Rd\mathbf{W}(s),
\end{aligned}
$$

where $\mathbf{c}(\mathbf{u}_n) = (c_1(\mathbf{u}_n), \ldots, c_d(\mathbf{u}_n)) \in \mathbb{R}^d$ with i-th component, as defined in (3.8).

By recurrence, we finally get the form of the approximate solution of (3.5) that is used for the convergence analysis.

$$
\mathbf{u}_n = e^{t_n A}\mathbf{u}_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - s)A}\mathbf{F}(\mathbf{u}_k)ds + (1-\alpha)\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+(1-\alpha)})A}G(\mathbf{u}_k)d\mathbf{W}(s) \tag{3.13}
$$

$$
\begin{aligned}
&+ \alpha\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\alpha})A}G(\mathbf{u}_k)d\mathbf{W}(s) + \alpha\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\alpha})A}S_3\mathbf{c}(\mathbf{u}_k)ds \\
&+ \alpha\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\alpha})A}DG(\mathbf{u}_k)S_2F(\mathbf{u}_k)d\mathbf{W}(s) \\
&+ \alpha\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\alpha})A}DG(\mathbf{u}_k)(S_1 - I)\mathbf{u}_k d\mathbf{W}(s) + \frac{\alpha}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\alpha})A}RdW(s).
\end{aligned}
$$

Before giving the form of the exact solution that we use for the proof of convergence, we require that the correction term $\mathbf{c}(\mathbf{u})$ satisfies the Lipschitz and the linear growth conditions of Assumption 2.1.2.

Now, let us examine the general SDE (3.5). Using Proposition 3.1.2 for interpreting the general stochastic integral of (3.9) by means of Itô calculus, the exact solution of (3.5) in integral form is equivalently written as

$$\mathbf{u}(t_n) = e^{t_nA}\mathbf{u}_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{F}(\mathbf{u}(s))ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}G(\mathbf{u}(s))d\mathbf{W}(s) \tag{3.14}$$

$$+\,\alpha\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s))ds,$$

where $\mathbf{c}(\mathbf{u})$ is as in (3.8) for $\phi(\mathbf{u}) = G(\mathbf{u})$. We consider the difference between the exact and the approximate solution given by (3.14)-(3.13) and we group it in different terms, in a convenient for analysis way. Similar to the Stratonovich approach of Section 2, we assume that there is no error in the initial data, which explains why there is no initial condition term among the difference terms defined below. Then, $\mathbf{u}(t_n) - \mathbf{u}_n = I + II + III + IV + V + VI$, where

$$I := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-s)A}\Big(\mathbf{F}(\mathbf{u}(s)) - \mathbf{F}(\mathbf{u}_k)\Big)ds$$

$$II := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \Big(e^{(t_n-s)A}G(\mathbf{u}(s)) - e^{t_nA}G(\mathbf{u}_k)\big((1-\alpha)e^{-(t_{k+(1-\alpha)})A} + \alpha e^{-t_{k+\alpha}A}\big)\Big)d\mathbf{W}(s)$$

$$III := \alpha\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \Big(e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s)) - e^{(t_n-t_{k+\alpha})A}S_3\mathbf{c}(\mathbf{u}_k)\Big)ds$$

$$IV := -\alpha\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\alpha})A}DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k)S_2 d\mathbf{W}(s)$$

$$V := -\alpha\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\alpha})A}DG(\mathbf{u}_k)(S_1 - I)\mathbf{u}_k d\mathbf{W}(s)$$

$$VI := -\frac{\alpha}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\alpha})A}R\,d\mathbf{W}(s).$$

Then, the strong or mean-square rate of convergence of the generalised version of the exponential integrators-based scheme of this section, is given by the following theorem.

Theorem 3.2.2. *Let $\boldsymbol{u}(t_n)$ denote the exact solution of (3.5) as given by equation (3.14) and $\boldsymbol{u}_n$ be the approximate solution given by equation (3.13). Moreover, let Assumptions 1.1.9, 1.1.10, 1.4.8, 2.1.2 and 3.2.1 hold. Then*

$$\Big(\mathbb{E}[\|\boldsymbol{u}(t_n) - \boldsymbol{u}_n\|_2^2]\Big)^{1/2} \leqslant C\Big(\Delta t^{1/2}\ln(\Delta t) + \Delta t^{p+\frac{1}{2}}\Big),$$

where $p > 0$ and C is a positive constant independent of Δt.

As before, we prove Theorem 3.2.2 by using a sequence of lemmas that provide separate upper bounds for each of the previously defined terms I-VI of the difference. Putting all the individual estimates together at the end, we deduce how well (3.13) approximates the exact solution.

Lemma 3.2.3.
Let Assumptions 1.1.9, 1.1.10 and 1.4.8 hold. Then

$$\mathbb{E}\big[\|I\|_2^2\big] \leqslant C_1 \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}[\|\boldsymbol{u}(s) - \boldsymbol{u}_k\|_2^2] ds,$$

where C_1 is a positive constant independent of Δt.

Proof. The proof is the same as in Lemma 2.1.3. □

Lemma 3.2.4.
Let Assumptions 1.1.9, 1.1.10 and 1.4.8 hold. Then

$$\mathbb{E}\big[\|II\|_2^2\big] \leqslant C_{2,1} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|\boldsymbol{u}(s) - \boldsymbol{u}_k\|_2^2\big] ds + C_{2,2} \Delta t \big(\ln(\Delta t)\big)^2,$$

where $C_{2,1}$ and $C_{2,2}$ are both positive constants independent of Δt.

Proof. By Itô's isometry

$$\mathbb{E}\big[\|II\|_2^2\big] = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|e^{(t_n-s)A} G(\mathbf{u}(s)) - e^{t_n A} G(\mathbf{u}_k)\big((1-\alpha)e^{-t_{k+(1-\alpha)}A} + \alpha e^{-t_{k+\alpha}A}\big)\|_F^2\big] ds.$$

By adding and subtracting the term $e^{(t_n-s)A}G(\mathbf{u}_k)$

$$\mathbb{E}\big[\|II\|_2^2\big] \leqslant \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|e^{(t_n-s)A}\big(G(\mathbf{u}(s)) - G(\mathbf{u}_k)\big)\|_F^2\big] ds$$

$$+ \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\|\Big(e^{(t_n-s)A} - e^{t_n A}\big((1-\alpha)e^{-t_{k+(1-\alpha)}A} + \alpha e^{-t_{k+\alpha}}\big)\Big) G(\mathbf{u}_k)\|_F^2\Big] ds.$$

That is,

$$\mathbb{E}\Big[\big\|II\big\|_2^2\Big] \leqslant II_1 + II_2,$$

where

$$II_1 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|e^{(t_n-s)A}\big(G(\mathbf{u}(s)) - G(\mathbf{u}_k)\big)\big\|_F^2\Big] ds$$

and

$$II_2 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\|\Big(e^{(t_n-s)A} - e^{t_nA}\big((1-\alpha)e^{-t_{k+(1-\alpha)}A} + \alpha e^{-t_{k+\alpha}}\big)\Big)G(\mathbf{u}_k)\|_F^2\Big]ds.$$

We seek upper bounds for the two different terms II_1 and II_2 separately. An upper bound for the term II_1 is given in (3.15) below and is found in a similar way to the bound for $\mathbb{E}\big[\|I\|_2^2\big]$.

$$II_1 \leqslant C_{2,1} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|\mathbf{u}(s) - \mathbf{u}_k\|_2^2\big]ds, \tag{3.15}$$

where $C_{2,1} = C_l^2 \sup_{0\leqslant s\leqslant t_n} \mathbb{E}[\|e^{(t_n-s)A}\|_F^2]$ and C_l is the Lipschitz constant as in Assumption 1.1.9. For the term II_2, we use Hölder's inequality, Remark 2.1.4 and the semigroup properties of Proposition 1.4.9.

$$II_2 = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\|\Big(e^{(t_n-s)A} - e^{(t_n-t_{k+(1-\alpha)})A} + \alpha(e^{(t_n-t_{k+(1-\alpha)})A} - e^{(t_n-t_{k+\alpha})A})\Big)G(\mathbf{u}_k)\|_F^2\Big]ds$$

$$\leqslant \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|G(\mathbf{u}_k)\|_F^2]$$

$$\times \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\|e^{(t_n-s)A} - e^{(t_n-t_{k+(1-\alpha)})A} + \alpha(e^{(t_n-t_{k+(1-\alpha)})A} - e^{(t_n-t_{k+\alpha})A})\|_F^2\Big]ds$$

$$\leqslant \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|G(\mathbf{u}_k)\|_F^2] \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\|e^{(t_n-s)A} - e^{(t_n-t_{k+\alpha})A}\|_F^2\Big]ds,$$

since $\alpha \in [0,1]$. Then

$$\begin{aligned}
II_2 \quad &\leqslant \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|G(\mathbf{u}_k)\|_F^2]C^2 \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \frac{(s-t_{k+\alpha})^2}{(t_n-s)^2}ds \\
&\leqslant \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|G(\mathbf{u}_k)\|_F^2]C^2 \sum_{k=0}^{n-1} \frac{1}{(t_n-t_{k+\alpha})^2} \int_{t_k}^{t_{k+1}} (s-t_{k+\alpha})^2 ds \\
&\leqslant \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|G(\mathbf{u}_k)\|_F^2]C^2 \Delta t \sum_{k=0}^{n-1} \left(\frac{\Delta t}{t_n - t_{k+\alpha}}\right)^2,
\end{aligned}$$

where C is a constant as in the semi-group properties Proposition 1.4.9. Then, since

$\sum_{k=0}^{n-1} \frac{\Delta t}{(t_n - t_{k-(1-\alpha)})}$ grows logarithmically, we have that for some constant $K > 0$

$$II_2 \leqslant \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|G(\mathbf{u}_k)\|_F^2]C^2K\Delta t\big(\ln(\Delta t)\big)^2 = C_{2,2}\Delta t\big(\ln(\Delta t)\big)^2,$$

where $C_{2,2} = \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|G(\mathbf{u}_k)\|_F^2]C^2K$. Putting the two estimates for II_1 and II_2 together finishes the proof.

□

Lemma 3.2.5.
Let Assumptions 1.1.9, 1.1.10, 1.4.8, 2.1.2 and 3.2.1 hold. Then

$$\mathbb{E}\big[\|III\|_2^2\big] \leqslant C_{3,1}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}[\|\boldsymbol{u}(s) - \boldsymbol{u}_k\|_2^2]ds + C_{3,2}\big(\Delta t(\ln(\Delta t))^2 + \Delta t^{2p+1}\big),$$

where $C_{3,1}$ and $C_{3,2}$ are positive constants independent of Δt and $p > 0$.

Proof.

$$\mathbb{E}[\|III\|_2^2] = \alpha\mathbb{E}\Big[\Big\|\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \big(e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s)) - e^{(t_n-t_{k+\alpha})A}S_3\mathbf{c}(\mathbf{u}_k)\big)ds\Big\|_2^2\Big].$$

By taking into consideration that $\alpha \leqslant 1$ and exchanging the order of expectation and integration

$$\mathbb{E}[\|III\|_2^2] \leqslant \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|\big(e^{(t_n-s)A}\mathbf{c}(\mathbf{u}(s)) - e^{(t_n-t_{k+\alpha})A}S_3\mathbf{c}(\mathbf{u}_k)\big)\|_2^2\big]ds.$$

We add and subtract the term $e^{(t_n-s)A}\mathbf{c}(\mathbf{u}_k)$ to get

$$\mathbb{E}\big[\|III\|_2^2\big] \leqslant III_1 + III_2$$

where

$$III_1 := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|e^{(t_n-s)A}\big(\mathbf{c}(\mathbf{u}(s)) - \mathbf{c}(\mathbf{u}_k)\big)\|_2^2\big]ds$$

and

$$III_2 := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|\big(e^{(t_n-s)A} - e^{(t_n-t_{k+\alpha})A}S_3\big)\mathbf{c}(\mathbf{u}_k)\|_2^2\big]ds.$$

Then, by Assumption 2.1.2 we have

$$III_1 \leqslant \sum_{k=0}^{n-1} \sup_{t_k \leqslant s \leqslant t_{k+1}} \mathbb{E}[\|e^{(t_n-s)A}\|_F^2] \int_{t_k}^{t_{k+1}} \mathbb{E}[\|\mathbf{c}(\mathbf{u}(s)) - \hat{\mathbf{c}}(\mathbf{u}_k)\|_2^2] ds$$

$$\leqslant C_{3,1} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}[\|\mathbf{u}(s) - \mathbf{u}_k\|_2^2] ds,$$

where $C_{3,1} = C_l^2 \sup_{0 \leqslant s \leqslant t_n} \mathbb{E}[\|e^{(t_n-s)A}\|_F^2]$ and C_l is the Lipschitz constant.

Moreover,

$$III_2 \leqslant \sup_{0 \leqslant k \leqslant n-1} \Big(\mathbb{E}[1 + \|\mathbf{c}(\mathbf{u}_k)\|_2^2]\Big) \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}[\|e^{(t_n-s)A} - e^{(t_n-t_{k+\alpha})A} S_3\|_F^2] ds.$$

At this stage, we additionally add and subtract $e^{(t_n-t_{k+\alpha})A}e^{\Delta t A}$ to the term $e^{(t_n-s)A} - e^{(t_n-t_{k+\alpha})A}S_3$ in order to create the difference $S - S_3$. Thus,

$$III_2 \leqslant \sup_{0 \leqslant k \leqslant n-1} \Big(\mathbb{E}[1 + \|\mathbf{c}(\mathbf{u}_k)\|_2^2]\Big)\big(III_{2,1} + III_{2,2}\big)$$

where

$$III_{2,1} := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}[\|e^{(t_n-s)A} - e^{(t_n-t_{k-(1-\alpha)})A}\|_F^2] ds$$

and

$$III_{2,2} := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}[\|e^{(t_n-t_{k+\alpha})A}(e^{\Delta t A} - S_3)\|_F^2] ds.$$

Using Proposition 1.4.9 and Assumption 3.2.1, we find upper bounds for the two new terms that have been formed.

$$III_{2,1} \leqslant C^2 \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \frac{(s - t_{k-(1-\alpha)})^2}{(t_n - s)^2} ds$$

$$\leqslant C^2 \sum_{k=0}^{n-1} \frac{1}{(t_n - t_{k-(1-\alpha)})^2} \int_{t_k}^{t_{k+1}} (s - t_{k-(1-\alpha)})^2 ds$$

$$\leqslant C^2 \Delta t \sum_{k=0}^{n-1} \left(\frac{\Delta t}{t_n - t_{k-(1-\alpha)}}\right)^2,$$

where C is a constant as in the semi-group properties Proposition 1.4.9. Then, since

$\sum_{k=0}^{n-1} \dfrac{\Delta t}{(t_n - t_{k-(1-\alpha)})}$ grows logarithmically, we have that for some constant $K > 0$

$$III_{2,1} \leqslant C^2 K \Delta t \big(\ln(\Delta t)\big)^2 = C_{3,21} \Delta t \big(\ln(\Delta t)\big)^2,$$

where $C_{3,21} = C^2 K$.

$$III_{2,2} \leqslant D^2 \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}[\|e^{(t_n - t_{k+\alpha})A}\|_F^2] \Delta t^{2p+1} = C_{3,22} \Delta t^{2p+1},$$

where $C_{3,22} = D^2 \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}[\|e^{(t_n - t_{k+\alpha})A}\|_F^2]$, $D, p > 0$ are constants as in Assumption 3.2.1.

Hence

$$III_2 \leqslant C_{3,2} \Big(\Delta t \big(\ln(\Delta t)\big)^2 + \Delta t^{2p+1}\Big),$$

with $C_{3,2} = \sup_{0 \leqslant k \leqslant n-1} \Big(\mathbb{E}[1 + \|\mathbf{c}(\mathbf{u}_k)\|_2^2]\Big) \max\{C_{3,21}, C_{3,22}\}$. Putting the estimates for III_1 and III_2 together, finishes the proof. □

Lemma 3.2.6. *Let Assumptions 1.1.9, 1.1.10, 1.4.8 and 3.2.1 hold. Then*

$$\mathbb{E}\big[\|IV\|_2^2\big] \leqslant C_4 \Delta t^{2q+1},$$

where C_4 is a positive constant independent of Δt and $q > 0$.

Proof. By Itô's isometry and Hölder's inequality

$$\mathbb{E}\big[\|IV\|_2^2\big] = \alpha \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|e^{(t_n - t_{k+\alpha})A} DG(\mathbf{u}_k) \mathbf{F}(\mathbf{u}_k) S_2\|_F^2\big] ds$$

$$\leqslant \alpha \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}[\|DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k)\|_F^2] \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}[\|e^{(t_n - t_{k+\alpha})A} S_2\|_F^2] ds.$$

Using Assumption 3.2.1,

$$\mathbb{E}\big[\|IV\|_2^2\big] \leqslant \alpha \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}[\|DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k)\|_F^2] \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}[\|e^{(t_n - t_{k+\alpha})A}\|_F^2] D \Delta t^{2q+1}$$

$$= C_4 \Delta t^{2q+1},$$

where $C_4 = \alpha D \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}[\|DG(\mathbf{u}_k)\mathbf{F}(\mathbf{u}_k)\|_F^2] \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}[\|e^{(t_n - t_{k+\alpha})A}\|_F^2]$ and $D > 0$ is a constant as in Assumption 3.2.1.

□

Lemma 3.2.7. *Let Assumptions 1.1.9-1.4.8 and 3.2.1 hold. Then*

$$\mathbb{E}\big[\|V\|_2^2\big] \leqslant C_5 \Delta t^{2p+1},$$

where $C_5 > 0$ is a constant independent of Δt and $p > 0$.

Proof. Again, by Itô's isometry and Hölder's inequality,

$$\mathbb{E}\big[\|V\|_2^2\big] = \alpha \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|e^{(t_n - t_{k+\alpha})A} DG(\mathbf{u}_k)(S_1 - I)\mathbf{u}_k\|_F^2\big] ds$$

$$\leqslant \alpha \sup_{0 \leqslant k \leqslant n-1} \mathbb{E}\big[\|DG(\mathbf{u}_k)\mathbf{u}_k\|_F^2\big] \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|e^{(t_n - t_{k+\alpha})A} S_1 - e^{(t_n - t_{k+\alpha})A}\|_F^2\big] ds.$$

By adding and subtracting the term $e^{(t_n - t_{k+\alpha})A}$

$$\mathbb{E}\big[\|V\|_2^2\big] \leqslant \alpha C \Big(\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|e^{(t_n - t_{k+\alpha})A}(S_1 - e^{\Delta t A})\|_F^2\big] ds$$

$$+ \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|e^{(t_n - t_{k-(1-\alpha)})A} - e^{(t_n - t_{k+\alpha})A}\|_F^2\big] ds \Big).$$

Using Assumption 3.2.1, we get the final estimate for this term. □

Lemma 3.2.8. *Let Assumptions 1.1.9, 1.1.10, 1.4.8, 2.1.1 and 3.2.1 hold and let $p > 0$. Then*

$$\mathbb{E}\big[\|VI\|_2^2\big] \leqslant C_6 \Delta t^{2p+1},$$

where C_6 is a positive constant independent of Δt.

Proof. By substituting the remainder R from equation (3.12)

$$\mathbb{E}[\|VI\|_2^2] = \frac{\alpha}{4} \mathbb{E}\Big[\| \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\alpha})A} \Big(\int_0^1 (1-\theta) \hat{\mathbf{h}}^\top D^2 G(\mathbf{u}_k + s\hat{\mathbf{h}})\hat{\mathbf{h}} d\theta \Big) d\mathbf{W}(s) \|_2^2 \Big].$$

Since, $(1-\theta)\hat{\mathbf{h}}^\top D^2 G(\mathbf{u}_k + \theta\hat{\mathbf{h}})\hat{\mathbf{h}}$ is bounded in the interval $[0,1]$, we have that

$$\int_0^1 (1-\theta)\hat{\mathbf{h}}^\top D^2 G(\mathbf{u}_k + \theta\hat{\mathbf{h}})\hat{\mathbf{h}} d\theta \leqslant \Delta t \max_{0 \leqslant \theta \leqslant 1} \big((1-\theta)\hat{\mathbf{h}}^\top D^2 G(\mathbf{u}_k + \theta\hat{\mathbf{h}})\hat{\mathbf{h}} \big).$$

Then,

$$\begin{aligned}
\mathbb{E}[\|VI\|_2^2] &\leqslant \frac{\alpha}{4}\mathbb{E}\Big[\|\sum_{k=0}^{n-1} e^{(t_n-t_{k+\alpha})A}\Delta t \max_{0\leqslant\theta\leqslant 1}\big((1-\theta)\hat{\mathbf{h}}^\top D^2G(\mathbf{u}_k+\theta\hat{\mathbf{h}})\hat{\mathbf{h}}\big)\Delta\mathbf{W}_k\|_2^2\Big] \\
&\leqslant \frac{\alpha}{4}\sup_{0\leqslant k\leqslant n-1}\Big\{\mathbb{E}[\|e^{(t_n-t_{k+\alpha})A}\|_F^2 \max_{0\leqslant\theta\leqslant 1}\big((1-\theta)\hat{\mathbf{h}}^\top D^2G(\mathbf{u}_k+\theta\hat{\mathbf{h}})\hat{\mathbf{h}}\big)]\Big\}\Delta t^3 \\
&= C_6\Delta t^3,
\end{aligned}$$

where $C_6 = \frac{\alpha}{4}\sup_{0\leqslant k\leqslant n-1}\Big\{\mathbb{E}[\|e^{(t_n-t_{k+\alpha})A}\|_F^2 \max_{0\leqslant\theta\leqslant 1}\big((1-\theta)\hat{\mathbf{h}}^\top D^2G(\mathbf{u}_k+s\hat{\mathbf{h}})\hat{\mathbf{h}}\big)]\Big\}$.

Alternatively, recall from the proof of Lemma 2.1.7 that $\max_{0\leqslant\theta\leqslant 1}\big((1-\theta)\hat{\mathbf{h}}^\top D^2G(\mathbf{u}_k+\theta\hat{\mathbf{h}})\hat{\mathbf{h}}\big)\Delta\mathbf{W}_k$ is a vector in $\mathbb{R}^d$, the i-th component of which is given by equation (2.15) after substituting $\hat{\mathbf{h}}$ for $\mathbf{h}$.

□

After estimating upper bounds for all the individual terms $I,\ldots,VI$, we can put the estimates together and prove the convergence Theorem 3.2.2.

Proof. Using the estimates provided by Lemmas 3.2.3, 3.2.4, 3.2.5 and 3.2.6, we get

$$\mathbb{E}[\|\mathbf{u}(t_n)-\mathbf{u}_n\|_2^2] \leqslant K\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}[\|\mathbf{u}(s)-\mathbf{u}_k\|_2^2]ds + L\Delta t\big(\ln(\Delta t)\big)^2 + M\Delta t^{2p+1},$$

where $K = C_1 + C_{2,1} + C_{3,1}$, $L = C_{2,2} + C_{3,2}$, $M = C_{3,2} + C_4 + C_5 + C_6$, all positive constants independent of Δt. Then, by Gronwall's lemma

$$\begin{aligned}
\mathbb{E}[\|\mathbf{u}(t_n)-\mathbf{u}_n\|_2^2] &\leqslant L\Delta t\big(\ln(\Delta t)\big)^2 + M\Delta t^{2p+1} \\
&\quad + K\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{K(t-s)}\Big(L\Delta t\big(\ln(\Delta t)\big)^2 + M\Delta t^{2p+1}\Big)ds \\
&\leqslant C^2\Big(\Delta t(\ln(\Delta t))^2 + \Delta t^{2p+1}\Big).
\end{aligned}$$

Taking the square root on both sides finishes the proof. □

In the above result for the order of convergence, the term $\Delta t^{p+\frac{1}{2}}$ is either smaller or equal to $\Delta t\big(\ln(\Delta t)\big)$ depending on what value the positive constant p takes. This confirms the observation made in [79] that as regards the overall accuracy, there is no particular advantage of using a predictor of higher order than the order of the corrector. It also means that $\Delta t^{p+\frac{1}{2}}$ could be ignored and hence omitted from the final rate of convergence result. However, we do not omit it since it reflects the choice of the method that we use for obtaining the predictor $\tilde{\mathbf{u}}_{n+1}$.

3.3 Numerical experiments on the Heston model

In this section, we experiment with the different values of α in scheme (3.10). We consider the one dimensional 3/2-Heston stochastic volatility model, used in finance for pricing derivatives. The model, as it appears in [25], is given by the following equation

$$du = \big(\mu u(t) - \kappa u^2(t)\big)dt + \lambda u(t)^{3/2} dW(t), \tag{3.16}$$

with $\kappa, \lambda, \mu > 0$. The authors in [25], introduce a Milstein-type scheme which they use for approximating (3.16). The scheme, called (θ, σ)-Milstein, is derived from the Stratonovich Taylor expansion and is given by the following equation

$$\begin{aligned} u_{n+1} = u_n &+ \big[\theta f(\tilde{u}_{n+1}) + (1-\theta) f(u_n)\big]\Delta t + g(u_n)\Delta W_n + \frac{1}{2} g'(u_n) g(u_n) \Delta W_n^2 \\ &- \frac{1}{2}\big[(1-\sigma) g'(u_n) g(u_n) + \sigma g'(\tilde{u}_{n+1}) g(\tilde{u}_{n+1})\big]\Delta t. \end{aligned} \tag{3.17}$$

With $\tilde{u}_{n+1}$ being given by the explicit expression

$$\begin{aligned} u_{n+1} &= \big(2(\kappa + \frac{3}{4}\lambda^2 \Delta t)\big)^{-1} \\ &\times \Big[\big((1-\mu\Delta t)^2 + 4(\kappa + \frac{3}{4}\lambda^2)\Delta t(u_n + \lambda u_n^{3/2}\Delta W_n + \frac{3}{4}\lambda^2 u_n^2 \Delta W_n^2)\big)^{1/2} - (1-\mu\Delta t)\Big]. \end{aligned} \tag{3.18}$$

In what follows, we approximate (3.16) by the scheme (3.17) and by our scheme (3.10) and we confirm that for $(\theta, \sigma) = (1, 1)$ and for $\alpha = 0.5$ we obtain the same results, see Figure 3.1. Later in Figure 3.2 we approximate (3.16) by (3.10) for different values of α corresponding to different interpretations of the stochastic integral. For $\alpha = 0$ we obtain the Itô interpretation while for $\alpha = \frac{1}{2}$ the Stratonovich.

In Figure 3.3, we look at the solution of (3.16) approximated by our scheme, as given in (3.10), for two different predictors obtained from the Euler-Maruyama and the ETD schemes.

Our scheme seems to perform well in approximating the solution of the Heston model considered at the specific time interval and for the chosen parameter values. However, note that the model does not satisfy the globally Lipschitz condition which is required for the strong convergence proof. This, could in general question the performance of the scheme for different parameter-values of the model, or when considered at a different time interval. A possible way to overcome this difficulty could be to consider a so called tamed scheme. Tamed schemes, as introduced in

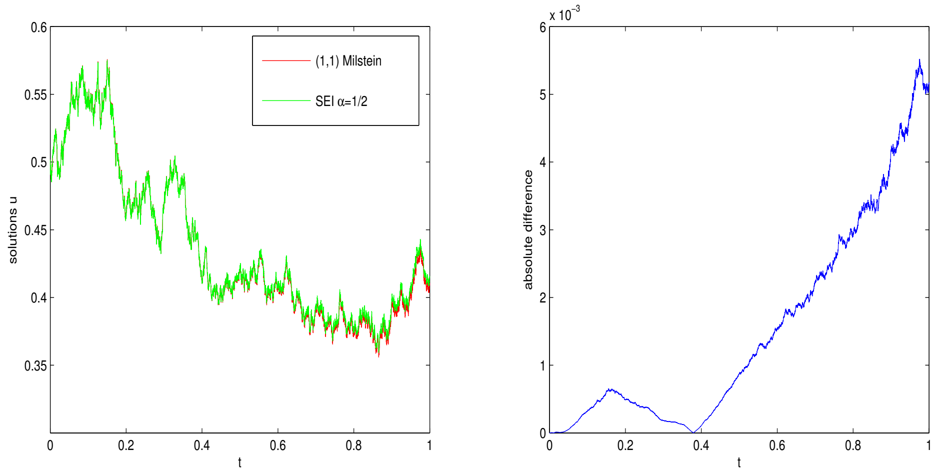


Figure 3.1: Heston stochastic volatility 3/2 model (3.16) approximated by (1, 1)-Milstein (3.17) and by general SEI (3.10) for $\alpha = 0.5$. Parameter values are $\mu = 0.1$, $\kappa = 0.2$, $\lambda = \sqrt{0.2}$ and initial condition $u(0) = 0.5$.

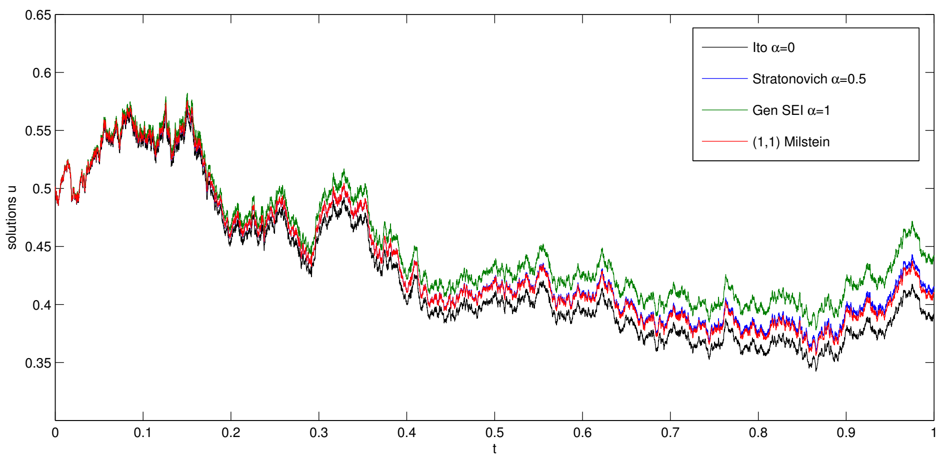


Figure 3.2: Heston stochastic volatility 3/2 model (3.16) approximated by (3.10) for different values $\alpha = 0, 0.5, 1$. Parameter values are $\mu = 0.1$, $\kappa = 0.2$, $\lambda = \sqrt{0.2}$ and initial condition $u(0) = 0.5$.

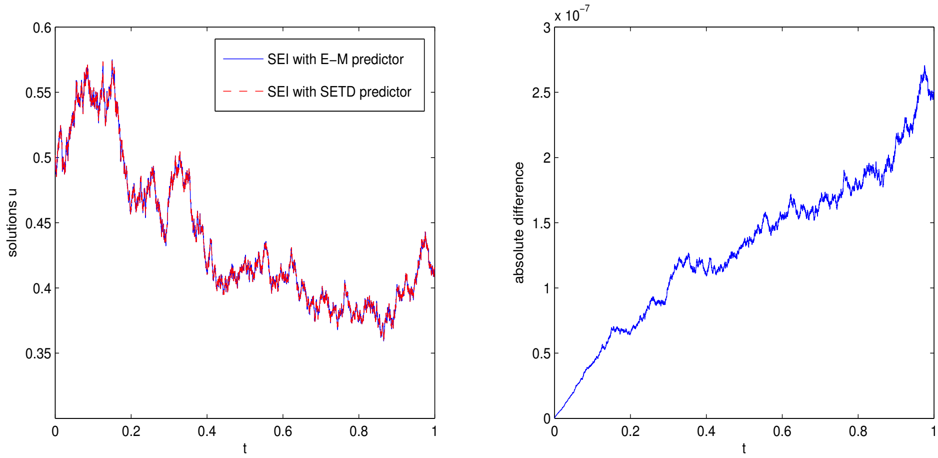


Figure 3.3: Heston stochastic volatility 3/2 model (3.16) approximated by (3.10) with $\alpha = 0.5$ and predictors $\tilde{u}_{n+1}$ taken by Euler-Maruyama and ETD schemes (left). Absolute difference between the two different solutions against time (right). Parameter values are $\mu = 0.1$, $\kappa = 0.2$, $\lambda = \sqrt{0.2}$ and initial condition $u(0) = 0.5$.

[31], comprise modifying the drift term in a specific way so that it is uniformly bounded. Another approach in [76], deals with problems where not only the drift but also the diffusion term is assumed to be only locally Lipschitz continuous. More results related to a class of suitably tamed numerical schemes can be found in [30].

This chapter concludes the SDEs section of this thesis. In the following chapters we deal with both the analysis and the numerical approximation of Stratonovich SPDEs.

Chapter 4

Stochastic PDEs

In this chapter, we extend the results of Chapter 2, from SDEs to a more general context where we are dealing with solving SPDEs. More particularly, we look at the approximation of SPDEs, using SEI as a time discretisation method and we prove analytically strong convergence. We mainly consider Stratonovich-type reaction-diffusion semilinear SPDEs of the form

$$du(x,t) = [Au(x,t) + F(u(x,t))]dt + G(u(x,t)) \circ dW(x,t). \tag{4.1}$$

For the sake of completeness however, we mention that a reaction-diffusion Itô-interpreted SPDE is given by

$$du(x,t) = [Au(x,t) + F(u(x,t))]dt + G(u(x,t))dW(x,t). \tag{4.2}$$

In this context, $u(x,t)$ of equation (4.1) is a random process taking values in a Hilbert space H. Note that $u(x,t)$ is suppressed notation for $u(x,t,\omega)$, where $\omega \in \Omega$ on a given filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbb{P})$ as seen in Definition 1.1.1. This will later be further suppressed to $u(t)$ for simplicity reasons. The initial condition $u(0,x)$ is denoted as $u^0(x)$. Our space and time variables lie in finite bounded domains D and $[0,T]$, $T > 0$ respectively. A is in general an unbounded linear operator that denotes the second spatial derivative ∂_{xx}, it acts on a Hilbert space H and is defined on a domain $\mathcal{D}(A) \subset H$. Unless stated otherwise H will be $L^2(D)$ in our case. We let A satisfy the following Assumption.

Assumption 4.0.1. *Let H be a Hilbert space and let A be a linear operator that generates an analytic semigroup $S(t) := e^{tA}$. We assume that there exists an orthonormal basis $\{\phi_j\}_{j\in\mathbb{N}}$ on H and also that A has the set of eigenfunctions $\{\phi_j\}_{j\in\mathbb{N}}$ with corresponding eigenvalues λ_j, $j \in \mathbb{N}$ in increasing order. Then, A can be rep-*

resented as

$$Au = \sum_{j=1}^{\infty} \lambda_j (\phi_j, u)\phi_j, \quad \textit{for all } u \in \mathcal{D}(A),$$

where the domain of A is defined as

$$\mathcal{D}(A) = \{u \in H : \sum_{j=1}^{\infty} |\lambda_j|^2 |(\phi_j, u)|^2 < \infty\}.$$

Then, A satisfies the properties of Proposition 1.4.8. Looking back at equation (4.1), we wish to give to $W(x,t)$ an interpretation that also involves the spatial variable and hence is suitable for the SPDEs context. The approach that we follow is that we consider $W(x,t)$ to be a space-time Q-Wiener process over a separable Hilbert space. By Q we denote the covariance operator of the Wiener process and we require that it satisfies the following assumption.

Assumption 4.0.2. *Let Q be a non-negative, symmetric, trace-class operator. Moreover, we assume that Q has an orthonormal basis $\{\phi_j\}_{j\in\mathbb{Z}}$ of eigenfunctions with corresponding eigenvalues ζ_j.*

Then, the trace of Q is defined by $Tr(Q) := \sum_{j\in\mathbb{Z}} \phi_j$. Note that, in order to have a trace class Q-Wiener process, we require that $Tr(Q) < +\infty$. In the opposite case, when $Q \equiv I$ for example, we have $Tr(Q) = +\infty$ and the process is called cylindrical Wiener process or space-time white noise [54]. We use the expansion of the following theorem to represent $W(t)$.

Theorem 4.0.3. *Let Q satisfy Assumption 4.0.2. Then $W(t)$ is a Q-Wiener process if and only if*

$$W(t) = \sum_{j\in\mathbb{Z}} \sqrt{\zeta_j}\phi_j\beta_j(t),$$

where $\beta_j(t)$ are independent, identically distributed (iid) Brownian motions and ζ_j, ϕ_j are the eigenvalues and the corresponding eigenfunctions of the covariance operator Q.

A proof of Theorem 4.0.3 can be found in [54]. Then, given a our spatial domain D, we choose $\{\phi\}_{j\in\mathbb{Z}}$ to be an orthonormal basis of $L^2(D)$ and we write $W(x,t)$ as

$$W(x,t) = \sum_{j\in\mathbb{Z}} \sqrt{\zeta_j}\phi_j(x)\beta_j(t). \tag{4.3}$$

$W(x,t)$ is now a Q-Wiener process that takes values in the Hilbert space $H^s(D)$, for $s \geqslant 0$. The choice of Q, or equivalently the choice of ϕ_j and ζ_j determines the type of noise that we have in space. In other words, there is a connection between the rate of decay of the eigenvalues ζ_j and the smoothness of the noise path in space. Note that, in order to simplify the notation, from now on we write $u(t)$ instead of $u(x,t)$ and $W(t)$ instead of $W(x,t)$.

Furthermore, we impose the following assumption on the drift term F.

Assumption 4.0.4. *The function $F : L^2(D) \to L^2(D)$ is continuous in $L^2(D)$ and satisfies*

1. *the Lipschitz condition*

$$\|F(u) - F(v)\|_{L^2(D)} \leqslant C_l \|u - v\|_{L^2(D)},$$

2. *the growth condition*

$$\|F(u)\|_{L^2(D)} \leqslant C_g \big(1 + \|u\|_{L^2(D)}\big),$$

for some constants $C_l, C_g > 0$ and for all $u, v \in L^2(D)$.

Before we impose the assumption on the diffusion term G, let us define the space where G takes values in.

Definition 4.0.5. *Let L_0^2 denote the space of linear operators $B : Q^{1/2}(H) \to H$ such that*

$$\|B\|_{L_0^2} := \left(\sum_{j=1}^{\infty} \|BQ^{1/2}\chi_j\|^2 \right)^{1/2} = \|BQ^{1/2}\|_{HS\left(Q^{1/2}(H),H\right)} < \infty$$

where $\{\chi_j\}_{j\in\mathbb{N}}$ is an orthonormal basis of H and $Q^{1/2}(H)$ is a subspace of the Hilbert space H with norm $\|u\|_{Q^{1/2}(H)} = \|Q^{-1/2}u\|$, $u \in H$.

The space $L_0^2 := HS\big(Q^{1/2}(H), H\big)$ is also called the space of Hilbert-Schmidt operators with corresponding norm $\|\cdot\|_{L_0^2}$ as defined above. Note that this definition is independent of the choice of the orthonormal basis. We are now ready to impose the following assumption on the diffusion term.

Assumption 4.0.6. *The function $G : L^2(D) \to L_0^2$, continuous in $L^2(D)$ satisfies*

1. *the Lipschitz condition*

$$\|G(u) - G(v)\|_{L_0^2} \leqslant C_l \|u - v\|_{L^2(D)},$$

2. *the growth condition*

$$\|G(u)\|_{L_0^2} \leqslant C_g \big(1 + \|u\|_{L^2(D)}\big),$$

for some constants $C_l, C_g > 0$ and for all $u, v \in L^2(D)$.

The above Assumptions ensure the existence and uniqueness of a mild solution for an Itô-interpreted SPDE. The following theorem summarises this result.

Theorem 4.0.7. *[54]*
Let A satisfy Assumption 1.4.8, $F : L^2(D) \to L^2(D)$ and $G : L^2(D) \to L^2_0$ satisfy Assumptions 4.0.4 and 4.0.6 respectively. If the initial condition $u_0 \in L^2(\Omega, \mathcal{F}_0, L^2(D))$, there exists a unique mild solution u to equation (4.2).

The mild solution u is satisfies the following equation

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A}F(u(s))ds + \int_0^t e^{(t-s)A}G(u(s))dW(s). \tag{4.4}$$

If we think of equation (4.1) though, as its equivalent corrected Itô SPDE with an additional term, as given by the Itô-Stratonovich correction formula, we can extend the mild solution result of [11, 13, 19, 54] to our context as well. We just need to impose the following extra assumption on the correction term $DG(u)G(u)$.

Assumption 4.0.8. *We assume that the function $G(u)DG(u)$ is continuous and satisfies*

1. *the Lipschitz condition*

$$\|G(u_1)DG(u_1) - G(u_2)DG(u_2)\|_{L^2_0} \leqslant L\|u_1 - u_2\|_{L^2(D)},$$

2. *the growth condition*

$$\|G(u)DG(u)\|_{L^2_0} \leqslant M\big(1 + \|u\|_{L^2(D)}\big),$$

for some constants $L, M > 0$ and for all $u_1, u_2 \in L^2(D)$. Note that $\|\cdot\|_{L^2_0}$ is the norm associated with the L^2_0 space of the Hilbert-Schmidt operators.

Note that a mild solution $u(t)$ satisfies the following inequalities for $p \geqslant 1$ and $\beta < 1$, see more in [41]

$$\sup_{0\leqslant t\leqslant T} \mathbb{E}[\|u(t)\|^p_{L^2(D)}] < \infty \tag{4.5}$$

and also

$$\sup_{0\leqslant t\leqslant T} \mathbb{E}[\|A^\beta u(t)\|^p_{L^2(D)}] < \infty. \tag{4.6}$$

We saw that based on Assumptions 1.4.8, 4.0.4, 4.0.6 and 4.0.8, Theorem 4.0.7 provides existence and uniqueness of a mild solution to equation (4.1). This is the so called semigroup approach [13]. Alternative approaches for the study of SPDEs are the variational approach [67, 47, 46] and the martingale approach [83]. There are also several either variational or semigroup approaches that allow the drift to be unbounded or random, see more for example in [2, 64, 69] and references therein.

After setting up the basics of the context for SPDEs, we are interested in approximating their solutions. In order to do that, we need to consider a discretization in both space and time. We have seen several methods for discretizing in time in Chapters 1, 2 and 3. In the following section, we make an introduction to space discretization.

4.1 Introduction to space discretization

There are several methods for discretising in space depending on the problem that we are solving, the domain in which we are solving it and the boundary conditions. In this chapter we will see how to solve parabolic SPDEs by considering finite differences (FD) and the Galerkin with finite elements method (FEM). As regards time discretisation, we will use the SEI scheme as displayed in Chapter 2.

4.1.1 Finite Difference (FD) method

One of the simplest ways to discretize in space is the FD method. The basic idea behind this method is to select a finite number of points, often called nodes, that create a mesh of our domain and then approximate the solution of the SPDE in these points. Practically, this means that the continuous in space and time solution is approximated by a discrete one, only known on the selected nodes. Of course, the finer the mesh, the more accurate the solution. There are several FD methods, one of the more standard of which being the uniformly spaced central differences.

We consider a discretisation on a finite grid $x_0, \ldots, x_m$, that is we have $m+1$ points, a spatial step Δx and we denote $u(x_j)$ by u_j. Then, the first spatial derivative ∂_x can be approximated by the central difference operator D_c given by

$$(D_c u)_j = \frac{u_{j+1} - u_{j-1}}{2\Delta x}, \quad j = 1, \ldots, m-1.$$

For the second spatial derivative, we define the operator $A_{\Delta x}$, which in matrix notation is given by

$$A_{\Delta x} = \begin{pmatrix} -2 & 2 & \\ & B & \\ & 2 & -2 \end{pmatrix} \in \mathbb{R}^{m \times m}, \text{ where } B = \begin{pmatrix} -2 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{pmatrix} \in \mathbb{R}^{(m-2)\times(m-2)},$$

and we approximate ∂_{xx} by $A_{\Delta x}$, also written as

$$(A_{\Delta x} u)_j = \frac{u_{j+1} - 2u_j + u_{j-1}}{\Delta x^2}. \tag{4.7}$$

As mentioned before and also as it is reflected in the definition of the differential operators, in our approach we consider Neumann boundary conditions. Hence, the discretised in space SPDE reads

$$du = [A_{\Delta x} u + F(u)]dt + G(u) \circ dW.$$

In order to obtain a fully discrete equation, we need to additionally discretise in time. For discretising in time we can use the SEI scheme as introduced in Chapter 2, an example of using finite differences and SEI is given later in Chapter 5.

4.1.2 Finite Element Method (FEM)

Another numerical method for discretizing in space is the FEM. FEM is a method, widely used in engineering design, that entails solving PDEs by dividing the domain of the problem into smaller bits, parts of a finite dimensional subspace of the original space in which we consider our problem. We briefly overview the idea in a deterministic context, in order to apply it later to the SPDEs case. There is extensive literature devoted to the construction, mathematical analysis and applications of the FEM. For further reading we refer to [8, 38, 51] and references therein.

Let us consider the inhomogeneous, one-dimensional parabolic (heat) equation with the additional term $f(t,x)$ (source/sink), over a domain $D \subset \mathbb{R}$, with smooth boundary ∂D, Dirichlet boundary conditions and initial condition $u(x,0) = u_0$

$$u_t = \alpha^2 u_{xx} + f(x,t), \ \text{ in } D$$

$$u = 0, \quad \text{on } \partial D. \tag{4.8}$$

Let us also define the space H_0^1 to be

$$H_0^1 = \{v : \|\nabla v\|_{L^2(D)} + \|v\|_{L^2(D)} < \infty \ \&\ v = 0 \text{ on } \partial D\}, \tag{4.9}$$

where $\|v\|_{L^2(D)}$ is the standard L^2-norm. Note that the definition of the above space depends on the boundary conditions that we have. We consider a partition of D given by $I_j = [x_j, x_{j+1}]$, $j = 1, \dots, m-1$ such that $\bigcup_{j=1}^{m-1} I_j = D_h$, where $h = x_{j+1} - x_j$.

We define the space $V_{h,0}$ to be

$$V_{h,0} = \{u_h \in C^0(\overline{D}) : u\Big|_{I_j} \text{ is linear } \& \ u = 0 \text{ outside } D_h\},$$

that is $u_h \in V_{h,0}$ are piece wise linear. Note that $V_{h,0}$ is finite dimensional and it is a subspace of H_0^1. Finally, let ϕ_j denote the hat functions given by

$$\phi_j(x) = \begin{cases} \frac{x-x_{j-1}}{x_j-x_{j-1}} & \text{if } x_{j-1} \leqslant x < x_j \\ \frac{x_{j+1}-x}{x_{j+1}-x_j} & \text{if } x_j \leqslant x < x_{j+1} \\ 0 & \text{otherwise} \end{cases}$$

Then, ϕ_j belong in $V_{h,0}$ for all j, they span the space $V_{h,0}$ and form a basis of H_0^1. The idea behind the Galerkin finite element method is to approximate the solution of the PDE by a function u_h which belongs in $V_{h,0}$.

After setting up this context, we can see how the FEM works by going back to equation (4.8). Let for example $D = (0, L), \alpha > 0$ and $t \in [0, T]$. Then, the variational formulation yields

$$\int_0^L u_t v dx + \int_0^L \alpha^2 u_x v_x dx = \int_0^L f \cdot v dx, \tag{4.10}$$

where u_t is called the trial and $v \in H_0^1$ is called test function. Moreover, v can be written as ϕ_i for some $i = 1, \ldots, m-1$ and a $u_h \in V_{h,0}$ can be written as $u_h = \sum_{j=1}^{N-1} \xi_j(t)\phi_j(x)$, where

$$\xi(t) = \begin{pmatrix} \xi_1(t) \\ \vdots \\ \xi_{m-1}(t) \end{pmatrix} = \begin{pmatrix} u_h(x_1, t) \\ \vdots \\ u_h(x_{m-1}, t) \end{pmatrix}.$$

Substituting v and u_h in (4.10) we get

$$\int_0^L f\phi_i(x)dx = \sum_{j=1}^{m-1} \xi_j'(t) \int_0^L \phi_j(x)\phi_i(x)dx + \sum_{j=1}^{m-1} \xi_j(t) \int_0^L \alpha^2 \phi_j'(x)\phi_i'(x)dx. \tag{4.11}$$

We denote

$$M_{ij} = \int_0^L \phi_j(x)\phi_i(x)dx,$$

$$A_{ij} = \int_0^L \alpha^2\phi_j'(x)\phi_i'(x),$$

$$b_i(t) = \int_0^L f\phi_i(x)dx$$

for $i, j = 1, \ldots, m-1$, called mass matrix, stiffness matrix and source respectively. Then (4.11) reads

$$b_i(t) = \sum_{j=1}^{m-1} \xi_j'(t)M_{ij} + \sum_{j=1}^{m-1} \xi_j(t)A_{ij}. \tag{4.12}$$

Alternatively, (4.12) can be written in matrix form

$$M\xi'(t) + A\xi(t) = b(t). \tag{4.13}$$

Solving (4.13) for $\xi_j(t)$, $j = 1, \ldots, m-1$,we get the approximate solution for the PDE, given by

$$u(x,t) \approx u_h(x,t) = \sum_{j=1}^{m-1} \xi_j(t)\phi_j(x).$$

4.2 FEM and SEI for reaction-diffusion SPDE

We can now combine spatial and time discretization in order to approximate the solution of (4.1). We use the FEM for discretizing in space and SEI for discretizing in time in this first approach.

Recall in equation (4.1) that W is a space-time Wiener process. Since we are interested in approximating the solution of (4.1), we also need to approximate the sample paths of the Q-Wiener process. Let us start from doing this by using the following finite sum

$$W_h(t) = \sum_{j=1}^{m-1} \sqrt{q_j}\phi_j(x)\beta_j(t). \tag{4.14}$$

In addition, we assume that $\{\phi_j\}_j$ span V_h and form an orthonormal basis of the space H_0^1. In other words, we assume that we can use the same set of basis functions $\{\phi_j\}_j$ to represent both the Wiener process $W(t)$ and $u(t)$. However, note that $\{\phi_j\}_j$ here is not necessarily the same as the basis of the operator Q in equation

(4.3). The strong solution of equation (4.1) is given by

$$u(t) = u_0 + \int_0^t [Au(s) + F(u(s))]ds + \int_0^t G(u(s)) \circ dW(s). \tag{4.15}$$

Variation of constants in (4.15) yields

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A}F(u(s))ds + \int_0^t e^{(t-s)A}G(u(s)) \circ dW(s). \tag{4.16}$$

Then, the weak formulation of the problem reads,

$$\langle u(t), v\rangle = \langle u_0, v\rangle + \int_0^t \big[\langle u(s), Av\rangle + \langle F(u(s)), v\rangle\big] ds + \int_0^t \langle G(u(s)) \circ dW(s), v\rangle, v \in V \tag{4.17}$$

where V is the space H_0^1 as defined by equation (4.9) and

$$\int_0^t \langle G(u(s)) \circ dW(s), v\rangle := \sum_{j=1}^{\infty} \int_0^t \langle G(u(s))\sqrt{q_j}\phi_j(x), v\rangle \circ d\beta_j(s).$$

Then, we are ready to discretize the SPDE in space by using the FEM. We introduce the finite dimensional subspace V_h and we are looking for an approximation $u_h \in V_h$ to the solution u. Hence, our problem becomes find u_h such that

$$\langle u_h(t), v\rangle = \langle u_{h,0}, v\rangle + \int_0^t \big[a(u_h(s), v) + \langle F(u_h(s)), v\rangle\big] ds + \langle \int_0^t G(u_h(s)) \circ dWv\rangle, \tag{4.18}$$

where $a(u_h, v) = \displaystyle\int_0^L \nabla u_h \nabla v dx = \langle u_h, v\rangle_{1/2}$, $v \in V_h$.

Let $D = [0, L]$ and let $P_h : L^2([0, L]) \to V_h$ be an orthogonal projection that projects the solution u to the finite dimensional space V_h. Also let $A_h : V_h \to V_h$ defined by $\langle A_h u, v\rangle = a(u, v)$ for all $u, v \in V_h$. We take some initial data $u_{h,0} = P_h u_0$ and from now on we follow the same approach as in [54]. We work on the finite-dimensional space V_h and we use the corresponding finite-dimensional version of the linear operator A, given by A_h. The approximate solution denoted by u_h satisfies the equation

$$du_h = \big[A_h u_h + P_h F(u_h)\big] dt + P_h G(u_h) \circ dW(t). \tag{4.19}$$

Note that $P_h G(u_h)$ lies in a finite dimensional space while $dW(t)$ does not. Hence, let $P_{N_W} : L^2(D) \to \text{span}\{\phi_1, \ldots, \phi_{N_W}\}$ such that

$$P_{N_W} W = W_h(t) = \sum_{j=1}^{N_W - 1} \sqrt{q_j}\phi_j(x)\beta_j(t),$$

that is P_{N_W} projects W in a finite dimensional space. Then, let $\mathcal{G}(s, u) := G(u)P_{N_W}$

and require that $\mathcal{G}$ satisfies the following assumption.

Assumption 4.2.1. *Let Assumption 4.0.6 hold. The function* $\mathcal{G} : \mathbb{R}^+ \times L^2(D) \to L_0^2$ *satisfies*

1.
$$\|\mathcal{G}(u_1) - \mathcal{G}(u_2)\|_{L_0^2} \leqslant L\|u_1 - u_2\|$$

2.
$$\|P_h\big(G(u(s)) - \mathcal{G}(u(t_k))\big)\|_{L^2(D,L_0^2)} \leqslant M\big(|s-t_k|^\theta + h^\zeta\big),$$

for some constants $L, M, \theta > 0$, *for all* $u_1, u_2 \in L^2(D)$ *and* $s > 0$.

Then, the semi-discrete equation reads

$$du_h = \big[A_h u_h + P_h F(u_h)\big]dt + P_h \mathcal{G}(u_h) \circ dW. \tag{4.20}$$

At this stage, we additionally discretize in time by applying to (4.20) the SEI scheme, introduced for SDEs in Chapter 2. So, the fully discretized equation is written as

$$u_{h,n+1} = e^{\Delta t A_h} u_{h,n} + A_h^{-1}(e^{\Delta t A_h} - I)P_h F(u_{h,n}) + \frac{1}{2} e^{\frac{\Delta t A_h}{2}} P_h\big(\mathcal{G}(u_{h,n}) + \mathcal{G}(\tilde{u}_{h,n+1})\big)\Delta W_n. \tag{4.21}$$

Recall that for deriving this scheme, we compute the deterministic integral and we approximate the stochastic one by using the definition of the Stratonovich integral in the following way

$$\int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A_h} P_h \mathcal{G}(u_{h,n}) \circ dW(s) \approx \frac{1}{2} e^{\frac{\Delta t A_h}{2}} P_h\big(\mathcal{G}(u_{h,n}) + \mathcal{G}(\tilde{u}_{h,n+1})\big)\Delta W_n.$$

Imitating the SDEs' case we can substitute for the predictor $\tilde{u}_{h,n+1}$ and Taylor-expand the term $\mathcal{G}(\tilde{u}_{h,n+1})$ so that

$$\mathcal{G}(u_{h,n} + \eta) = \mathcal{G}(u_{h,n}) + D\mathcal{G}(u_{h,n})\eta + R_1$$

where $\eta := \big(A_h u_{h,n} + P_h F(u_{h,n})\big)\Delta t + \mathcal{G}(u_{h,n})\Delta W_n$. In this approach, $\tilde{u}_{h,n+1}$ is obtained by the Euler-Maruyama method, however, there are alternative methods that can be used in order to obtain $\tilde{u}_{h,n+1}$. Recall from the SDEs proof for the generalised version of the predictor that the order of convergence for the method that we use, contributes to the final order of convergence for the scheme, as seen in Theorem 3.2.2. Hence (4.21) can be rewritten as

$$u_{h,n+1} = e^{\Delta t A_h} u_{h,n} + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A_h} P_h F(u_{h,n}) ds + \int_{t_n}^{t_{n+1}} e^{\frac{\Delta t A_h}{2}} P_h \mathcal{G}(u_{h,n}) dW(s)$$

$$+ \frac{1}{2} \int_{t_n}^{t_{n+1}} e^{\frac{\Delta t A_h}{2}} P_h D\mathcal{G}(u_{h,n}) \eta dW(s) + \frac{1}{4} \int_{t_n}^{t_{n+1}} e^{\frac{\Delta t A_h}{2}} P_h R_1 dW(s).$$

Expanding η and considering the equation piece-wisely in the time interval $[0, t_n]$ leads us to the following recurrence, which in fact is the most appropriate form of the approximate solution to use for the analysis.

$$u_{h,n} = e^{t_n A_h} u_{h,0} + \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n-s)A_h} P_h F(u_{h,k}) ds \tag{4.22}$$

$$+ \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h \mathcal{G}(u_{h,k}) dW(s)$$

$$+ \frac{1}{2} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h \mathcal{G}(u_{h,k}) D\mathcal{G}(u_{h,k}) ds$$

$$+ \frac{1}{2} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h D\mathcal{G}(u_{h,k}) F(u_{h,k}) \Delta t dW(s)$$

$$+ \frac{1}{4} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h \Big(\int_0^1 (1-s) \eta^\top D^2 \mathcal{G}(u_{h,k} + s\eta) \eta ds \Big) dW(s).$$

Recall from Theorem 4.0.7 that assumptions 4.0.1, 4.0.4, 4.0.6 provide the necessary requirements for the existence of a mild solution. The following lemma creates the necessary context for regularity of the mild solution in time.

Lemma 4.2.2. *Let A, F and G satisfy Assumptions 4.0.1, 4.0.4 and 4.0.6 respectively. For $T > 0$, $\epsilon \in (0, \zeta)$ and $\theta_1 = \min\{(\zeta - \epsilon)/2, 1/2\}$ there exists $K_{RT} > 0$ such that*

$$\|u(t_2) - u(t_1)\|_{L^2(D, L^2(D))} \leqslant K_{RT} (t_2 - t_1)^{\theta_1}, \quad 0 \leqslant t_1 \leqslant t_2 \leqslant T.$$

Proof. The proof for the regularity of the solution in time can be found in [54]. □

Hence, the exact solution of (4.1) after using the Itô-Stratonovich correction formula, the variation of constants and considering it in each of the subintervals

$[t_k, t_{k+1})$ for $k = 0, \dots, n-1$ satisfies the following equation

$$u(t) = e^{t_n A}u_0 + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - s)A}F(u(s))ds + \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - s)A}G(u(s))dW(s) \tag{4.23}$$

$$+ \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - s)A}G(u(s))DG(u(s))Tr(Q)ds,$$

where $Tr(Q)$ stands for the trace of the covariance operator Q.

We are interested in proving that the approximate solution, which is discretised in space by using the FEM and in time by means of SEI, converges to the exact one. Following almost the same steps with the proof of convergence for SDEs, we are going to consider the difference between the exact and the approximate solution given by (4.23)-(4.22). The difference can be grouped in five different terms for which we find upper bound estimates. These estimates show that the approximation converges and determine the order of convergence as well.

$$I := e^{t_n A}u_0 - e^{t_n A_h}P_h u_{h,0}$$

$$II := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \Big[e^{(t_n - s)A}F(u(s)) - e^{(t_n - s)A_h}P_h F(u_{h,k})\Big]ds$$

$$III := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \Big[e^{(t_n - s)A}G(u(s)) - e^{(t_n - t_{k+\frac{1}{2}})A_h}P_h \mathcal{G}(u_{h,k})\Big]dW(s)$$

$$IV := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \Big[e^{(t_n - s)A}G(u(s))DG(u(s))Tr(Q) - e^{(t_n - t_{k+\frac{1}{2}})A_h}P_h \mathcal{G}(u_{h,k})D\mathcal{G}(u_{h,k})\Big]ds$$

$$V := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\frac{1}{2}})A_h}P_h D\mathcal{G}(u_{h,k})\mathcal{G}(u_{h,k})\Delta t dW(s)$$

$$+ \frac{1}{4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\frac{1}{2}})A_h}P_h\Big(\int_0^1 (1-s)\eta^\top D^2\mathcal{G}(u_{h,k} + s\eta)\eta ds\Big)dW(s).$$

At this stage we define the operator $T_{h,n}$ by $T_{h,n}u_0 := (e^{t_n A} - e^{t_n A_h}P_h)u_0$. The following, useful for our analysis, estimate holds.

Lemma 4.2.3. *[81] Let $T_{h,n}$ be an operator defined as above such that $T_{h,n}u_0 =$*

$u(t_n) - u_{h,n}$. *Then*

$$\|T_{h,n}u_0\|_{L^2(D)} \leqslant K\frac{\Delta t + h^{2\beta}\Delta t^{1-\beta} + h^2}{t_n^{1-\beta}}\|u_0\|_\beta, \quad \textit{for all } u \in \mathcal{D}(A^\beta)$$

for some constant $K > 0$.

The following lemmas provide error estimates for all the above difference terms. Putting these estimates together at the end gives the final estimate.

Lemma 4.2.4. *Let Assumption 4.0.1 hold, then there exists a constant* C_1 *such that*

$$\mathbb{E}[\|I\|^2_{L^2(D)}] \leqslant C_1(\Delta t + h^2)^2.$$

Proof.

$$\mathbb{E}[\|I\|^2_{L^2(D)}] = \mathbb{E}\big[\|(e^{t_nA} - e^{t_nA_h}P_h)u_0\|^2_{L^2(D)}\big].$$

By the definition of $T_{h,n}$ and Lemma 4.2.3 with $\beta = 1$, we have

$$\mathbb{E}[\|I\|^2_{L^2(D)}] \leqslant K^2(\Delta t + 2h^2)^2\|u_0\|_{L^2(D)},$$

where K is a positive constant as in Lemma 4.2.3. □

Lemma 4.2.5. *Let Assumptions 4.0.1, 4.0.4 hold, then*

$$\mathbb{E}[\|II\|^2_{L^2(D)}] \leqslant C_{2,1}\Delta t\big(\ln(\Delta t)\big)^2 + C_{2,2}\ln(\Delta t)(\Delta t + h^2)^2$$

$$+ C_{2,3}\Delta t^{2(1-\epsilon)} + C_{2,4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\big[\|u(t_k) - u_{h,k}\|^2_{L^2(D)}\big]ds.$$

Proof. Let us start from observing that $II = II_1 + II_2 + II_3 + II_4$ where

$$II_1 := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\big(e^{(t_n-s)A}F(u(s)) - e^{(t_n-t_k)A}F(u(s))\big)ds$$

$$II_2 := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\big(e^{(t_n-t_k)A}F(u(s)) - e^{(t_n-t_k)A_h}P_hF(u(s))\big)ds$$

$$II_3 := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\big(e^{(t_n-t_k)A_h}P_hF(u(s)) - e^{(t_n-t_k)A_h}P_hF(u(t_k))\big)ds$$

$$II_4 := \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\big(e^{(t_n-t_k)A_h}P_hF(u(t_k)) - e^{(t_n-t_k)A_h}P_hF(u_{h,k})\big)ds.$$

By the growth condition satisfied by the drift term as given in Assumption 4.0.4 and by Proposition 1.4.8 we have

$$\begin{aligned}
\mathbb{E}[\|II_1\|^2_{L^2(D)}] &\leqslant \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|\big(e^{(t_n-s)A}-e^{(t_n-t_k)A}\big)F(u(s))\big\|^2_{L^2(D)}\Big]ds \\
&\leqslant M^2 \sup_{0\leqslant s\leqslant t_n}\big(1+\|u(s)\|^2_{L^2(D)}\big)\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\mathbb{E}[\|e^{(t_n-s)A}-e^{(t_n-t_k)A}\|^2_{L^2_0}]ds. \\
&\leqslant M^2 \sup_{0\leqslant s\leqslant t_n}\big(1+\|u(s)\|^2_{L^2(D)}\big)\left(\Delta t^2+\sum_{k=0}^{n-2}\int_{t_k}^{t_{k+1}}K^2\Big(\frac{s-t_k}{t_n-s}\Big)^2 ds\right) \\
&\leqslant M^2K^2 \sup_{0\leqslant s\leqslant t_n}\big(1+\|u(s)\|^2_{L^2(D)}\big)\Big(\Delta t^2+\sum_{k=0}^{n-2}\frac{1}{(t_n-t_{k+1})^2}\int_{t_k}^{t_{k+1}}(s-t_k)^2ds\Big)
\end{aligned}$$

$$\begin{aligned}
&\leqslant M^2K^2 \sup_{0\leqslant s\leqslant t_n}\big(1+\|u(s)\|^2_{L^2(D)}\big)\left(\Delta t^2+\Delta t\sum_{k=0}^{n-1}\Big(\frac{\Delta t}{t_n-t_{k+1}}\Big)^2\right) \\
&\leqslant C_{2,1}\Delta t\big(\ln(\Delta t)\big)^2,
\end{aligned}$$

where $C_{2,1} = 2M^2K^2 \sup_{0\leqslant s\leqslant t_n}\big(1+\|u(s)\|^2_{L^2(D)}\big)$ and $M, K > 0$ are constants as in Assumption 4.0.6 and Proposition 1.4.8 respectively. For the next term we use the definition of $T_{h,n}$ together with the corresponding Lemma 4.2.3 and Assumption 4.0.4 to get

$$\begin{aligned}
\mathbb{E}[\|II_2\|^2_{L^2(D)}] &\leqslant \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|\big(e^{(t_n-t_k)A}-e^{(t_n-t_k)A_h}P_h\big)F(u(s))\big\|^2_{L^2(D)}\Big]ds \\
&= \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|T_{h,n-k}F(u(s))\|^2_{L^2(D)}\big]ds \\
&\leqslant \sup_{0\leqslant s\leqslant t_n}\big(1+\|u(s)\|^2_{L^2(D)}\big)\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}K^2\frac{(\Delta t+h^2)^2}{t_n-t_{k+\frac{1}{2}}}ds \\
&\leqslant C_{2,2}\ln(\Delta t)(\Delta t+h^2)^2,
\end{aligned}$$

where $C_{2,2} = K^2 \sup_{0\leqslant s\leqslant t_n}\big(1+\|u(s)\|^2_{L^2(D)}\big)$ and K is a positive constant as in Proposition 1.4.8.

For the third term, we use the boundedness of the operator $S(t)$, the Lipschitz condition from Assumption 4.0.4 and Lemma 4.2.2,

$$
\begin{aligned}
\mathbb{E}[\|II_3\|^2_{L^2(D)}] &\leqslant \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|e^{(t_n-t_k)A_h}P_h\big(F(u(s))-F(u(t_k))\big)\big\|^2_{L^2(D)}\Big]ds \\
&\leqslant L^2 \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|e^{(t_n-t_k)A_h}P_h\|^2_{L^2_0}]\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(s))-u(t_k)\|^2_{L^2(D)}\big]ds \\
&\leqslant L^2 \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|e^{(t_n-t_k)A_h}P_h\|^2_{L^2_0}]\sum_{k=0}^{n-1}(s-t_k)^{2(1-\epsilon)} \leqslant C_{2,3}\Delta t^{2(1-\epsilon)}
\end{aligned}
$$

where $C_{2,3} = L^2 \sup\limits_{0\leqslant k\leqslant n-1} \mathbb{E}[\|e^{(t_n-t_k)A_h}P_h\|^2_{L^2_0}]$ and L is the Lipschitz constant. Using again the Lipschitz condition from Assumption 4.0.4, we bound the final term in a similar way to II_3.

$$
\begin{aligned}
\mathbb{E}[\|II_4\|^2_{L^2(D)}] &\leqslant \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|e^{(t_n-t_k)A_h}P_h\big(F(u(t_k))-F(u_{h,k})\big)\big\|^2_{L^2(D)}\Big]ds \\
&\leqslant L^2 \sup_{0\leqslant k\leqslant n-1} \mathbb{E}[\|e^{(t_n-t_k)A_h}P_h\|^2_{L^2_0}]\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(t_k)-u_{h,k}\|^2_{L^2(D)}\big]ds \\
&= C_{2,4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(t_k)-u_{h,k}\|^2_{L^2(D)}\big]ds,
\end{aligned}
$$

where $C_{2,4} = L^2 \sup\limits_{0\leqslant k\leqslant n-1} \mathbb{E}[\|e^{(t_n-t_k)A_h}P_h\|^2_{L^2_0}]$ and L is the Lipschitz constant. Putting all the individual estimates for II_1, II_2, II_3 and II_4 together, gives the overall estimate for the term $\mathbb{E}[\|II\|^2_{L^2(D)}]$. □

Lemma 4.2.6. *Let Assumptions 4.0.1, 4.0.6 hold, then*

$$
\begin{aligned}
\mathbb{E}\big[\|III\|^2_{L^2(D)}\big] &\leqslant C_{3,1}\Delta t\big(\ln(\Delta t)\big)^2 + C_{3,2}\ln(\Delta t)(\Delta t+h^2)^2 \\
&\quad + C_{3,3}(\Delta t^{2\theta}+h^{2\zeta}) + C_{3,4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(t_k)-u_{h,k}\|^2_{L^2(D)}\big]ds.
\end{aligned}
$$

Proof. Working in a similar way with the previous lemma, we split this term in four different terms and we work separately on each term. More specifically,

$$III = III_1 + III_2 + III_3 + III_4$$

where

$$III_1 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \big(e^{(t_n-s)A} G(u(s)) - e^{(t_n - t_{k+\frac{1}{2}})A} G(u(s)) \big) dW(s)$$

$$III_2 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \big(e^{(t_n - t_{k+\frac{1}{2}})A} G(u(s)) - e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h G(u(s)) \big) dW(s)$$

$$III_3 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \big(e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h G(u(s)) - e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h \mathcal{G}(u(t_k)) \big) dW(s)$$

$$III_4 := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \big(e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h \mathcal{G}(u(t_k)) - e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h \mathcal{G}(u_{h,k}) \big) dW(s).$$

For the first term, by Itô's isometry, Assumption 4.0.6 for the diffusion term G and Proposition 1.4.8 for $\alpha = \beta = 1$ we get

$$\mathbb{E}[\|III_1\|^2_{L^2(D)}] = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\| \big(e^{(t_n-s)A} - e^{(t_n - t_{k+\frac{1}{2}})A} \big) G(u(s)) \big\|^2_{L^2_0} \Big] ds$$

$$\leqslant M^2 \sup_{0 \leqslant s \leqslant t_n} \big(1 + \|u(s)\|^2_{L^2(D)}\big) K^2 \Big(\Delta t^2 + \sum_{k=0}^{n-2} \frac{1}{(t_n - t_{k+1})^2} \int_{t_k}^{t_{k+1}} (s - t_{k+\frac{1}{2}})^2 ds \Big)$$

$$\leqslant M^2 K^2 \sup_{0 \leqslant s \leqslant t_n} \big(1 + \|u(s)\|^2_{L^2(D)}\big) \Big(\Delta t^2 + \Delta t \big(\ln(\Delta t) \big)^2 \Big)$$

$$\leqslant C_{3,1} \Delta t \big(\ln(\Delta t) \big)^2,$$

where $C_{3,1} = 2M^2K^2 \sup_{0 \leqslant s \leqslant t_n} \big(1 + \|u(s)\|^2_{L^2(D)}\big)$ and $M, K > 0$ are constants as in Assumption 4.0.6 and Proposition 1.4.8 respectively. For the second term, we use again Itô's isometry, Lemma 4.2.3 that provides an estimate for $T_{h,n-(k+\frac{1}{2})}$ and Assumption 4.0.6.

$$
\begin{aligned}
\mathbb{E}[\|III_2\|^2_{L^2(D)}] &= \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|\big(e^{(t_n-t_{k+\frac{1}{2}})A} - e^{(t_n-t_{k+\frac{1}{2}})A_h}P_h\big)G(u(s))\big\|^2_{L^2_0}\Big]ds \\
&= \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|T_{h,n-(k+\frac{1}{2})}G(u(s))\|^2_{L^2_0}\big]ds \\
&\leqslant K^2 \sup_{0\leqslant s\leqslant t_n}\big(1+\|u(s)\|^2_{L^2(D)}\big)\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \frac{(\Delta t+h^2)^2}{(t_n-t_{k+\frac{1}{2}})}ds \\
&\leqslant C_{3,2}\ln(\Delta t)(\Delta t+h^2)^2.
\end{aligned}
$$

Note that, $\frac{1}{t_n-t_{k+\frac{1}{2}}} > 1$ as k varies from 0 to $n-1$. For the third individual term, we use Assumption 4.2.1 and the boundedness of the operator P_h, to get

$$
\begin{aligned}
\mathbb{E}[\|III_3\|^2_{L^2(D)}] &= \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|e^{(t_n-t_{k+\frac{1}{2}})A_h}P_h\big(G(u(s))-\mathcal{G}(u(t_k))\big)\big\|^2_{L^2_0}\Big]ds \\
&\leqslant CM^2\big(|s-t_k|^{2\theta}+h^{2\zeta}\big) \\
&\leqslant C_{3,3}\big(\Delta t^{2\theta}+h^{2\zeta}\big),
\end{aligned}
$$

since s varies between t_k and t_{k+1}. Moreover, $C_{3,3}$ is a constant that contains $C, M > 0$, also constants associated with the the boundedness of P_h and with Assumption 4.2.1 respectively. Finally, for the fourth term, by Proposition 1.4.8 and Assumption 4.2.1 we have

$$
\begin{aligned}
\mathbb{E}[\|III_4\|^2_{L^2(D)}] &= \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|e^{(t_n-t_{k+\frac{1}{2}})A_h}P_h\big(\mathcal{G}(u(t_k))-\mathcal{G}(u_{h,k})\big)\big\|^2_{L^2_0}\Big]ds \\
&\leqslant C_{3,4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(t_k)-u_{h,k}\|^2_{L^2(D)}\big]ds.
\end{aligned}
$$

Putting all four estimates together completes the proof. □

Lemma 4.2.7. *Let Assumptions 4.0.1, 4.0.8 hold, then*

$$\mathbb{E}\big[\|IV\|^2_{L^2(D)}\big] \leqslant C_{4,1}\Delta t\big(\ln(\Delta t)\big)^2 + C_{4,2}\ln(\Delta t)(\Delta t + h^2)^2$$

$$+ C_{4,3}\Delta t^{2(1-\epsilon)} + C_{4,4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(t_k) - u_{h,k}\|^2_{L^2(D)}\big]ds.$$

Proof. This term is also split into four different terms so that

$$IV = IV_1 + IV_2 + IV_3 + IV_4,$$

where

$$IV_1 := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \big(e^{(t_n-s)A}G(u(s))DG(u(s)) - e^{(t_n-t_{k+\frac{1}{2}})A}G(u(s))DG(u(s))\big)ds$$

$$IV_2 := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n-t_{k+\frac{1}{2}})A}G(u(s))DG(u(s)) - e^{(t_n-t_{k+\frac{1}{2}})A_h}P_hG(u(s))DG(u(s))\big)ds$$

$$IV_3 := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \big(e^{(t_n-t_{k+\frac{1}{2}})A_h}P_hG(u(s))DG(u(s))$$

$$- e^{(t_n-t_{k+\frac{1}{2}})A_h}P_hG(u(t_k))DG(u(t_k))\big)ds$$

$$IV_4 := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \big(e^{(t_n-t_{k+\frac{1}{2}})A_h}P_hG(u(t_k))DG(u(t_k))$$

$$- e^{(t_n-t_{k+\frac{1}{2}})A_h}P_hG(u_{h,k})DG(u_{h,k})Tr(Q)\big)ds.$$

We handle IV_1 in a similar way to III_1. By Assumption 4.0.8 and Proposition 1.4.8 for $\alpha = \beta = 1$ we get

$$\mathbb{E}[\|IV_1\|^2_{L^2(D)}] \leqslant \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\Big\|\big(e^{(t_n-s)A} - e^{(t_n-t_{k+\frac{1}{2}})A}\big)G(u(s))DG(u(s))\Big\|^2_{L^2_0}\Big]ds$$

$$\leqslant M^2K^2\sup_{0\leqslant s\leqslant t_n}\big(1 + \|u(s)\|^2_{L^2(D)}\big)\Big(\Delta t^2 + \Delta t\big(\ln(\Delta t)\big)^2\Big)$$

$$\leqslant C_{4,1}\Delta t\big(\ln(\Delta t)\big)^2,$$

where $C_{4,1} = 2M^2K^2 \sup_{0\leqslant s\leqslant t_n} \big(1+\|u(s)\|^2_{L^2(D)}\big)$ and $M, K > 0$ are constants as in Assumption 4.0.8 and Proposition 1.4.8 respectively.

For the second term, by the definition of $T_{h,n-(k+\frac{1}{2})}$ and Lemma 4.2.3 we have that

$$\begin{aligned}
\mathbb{E}[\|IV_2\|^2_{L^2(D)}] &\leqslant \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|\big(e^{(t_n-t_{k+\frac{1}{2}})A} - e^{(t_n-t_{k+\frac{1}{2}})A_h}\big)G(u(s))DG(u(s))\big\|^2_{L^2_0}\Big]ds \\
&= \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|T_{h,n-(k+\frac{1}{2})}G(u(s))DG(u(s))\|^2_{L^2_0}\big]ds \\
&\leqslant C_{4,2}\ln(\Delta t)(\Delta t + h^2)^2,
\end{aligned}$$

where $C_{4,2} = \frac{1}{2}M^2 \sup_{0\leqslant s\leqslant t_n} \big(1+\|u(s)\|^2_{L^2(D)}\big)$ and $M > 0$ is a constant as in Assumption 4.0.8. Moreover, by the boundedness of P_h and by the time-regularity condition satisfied by the solution, as seen in Lemma 4.2.2, we get

$$\mathbb{E}[\|IV_3\|^2_{L^2(D)}] \leqslant \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|e^{(t_n-t_{k+\frac{1}{2}})A_h}P_h\big(G(u(s))DG(u(s)) - G(u(t_k))DG(u(t_k))\big)\big\|^2_{L^2_0}\Big]ds$$

$$\begin{aligned}
&\leqslant C\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(s)-u(t_k)\|^2_{L^2(D)}\big]ds \\
&\leqslant C\sum_{k=0}^{n-1}(s-t_k)^{2(1-\epsilon)} \\
&\leqslant C_{4,3}\Delta t^{2(1-\epsilon)}.
\end{aligned}$$

Finally,

$$\begin{aligned}
\mathbb{E}[\|IV_4\|^2_{L^2(D)}] \leqslant \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} &\mathbb{E}\Big[\big\|e^{(t_n-t_{k+\frac{1}{2}})A_h}Tr(Q)P_h \\
&\times\big(G(u(t_k))DG(u(t_k)) - G(u_{h,k})DG(u_{h,k})\big)\big\|^2_{L^2_0}\Big]ds
\end{aligned}$$

$$\mathbb{E}[\|IV_4\|^2_{L^2(D)}] \leqslant CTr(Q)\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(t_k) - u_{h,k}\|^2_{L^2(D)}\big]ds$$

$$= C_{4,4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(t_k) - u_{h,k}\|^2_{L^2(D)}\big]ds.$$

Putting the individual estimates together, clearly gives the result. □

Finally, in the following lemma we are left with showing that the term V not only is bounded but also significantly smaller than the previously examined terms.

Lemma 4.2.8. *The term $\mathbb{E}[\|V\|^2]$ is bounded in the following way*

$$\mathbb{E}[\|V\|^2_{L^2(D)}] \leqslant C_{5,1}\Delta t^3 + C_{5,2}(\Delta t^{7/2} + \Delta t^3 + \Delta t^{5/2}).$$

Proof. In a similar way to the proofs of the previous lemmas, let $V = V_1 + V_2$ where

$$V_1 := \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h D\mathcal{G}(u_{h,k})F(u_{h,k})\Delta t dW(s)$$

$$V_2 := \frac{1}{4}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t - t_{k+\frac{1}{2}})A_h} P_h\Big(\int_0^1 (1-s)\eta^\top D^2\mathcal{G}(u_{h,k} + s\eta)\eta ds\Big)dW(s).$$

The upper bound for the term V_1 is found by using Itô's isometry and the boundedness of P_h,

$$\mathbb{E}[\|V_1\|^2_{L^2(D)}] = \frac{1}{2}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\Big\|e^{(t_n - t_{k+\frac{1}{2}})A_h} P_h D\mathcal{G}(u_{h,k})F(u_{h,k})\Delta t\Big\|^2_{L^2(D)}\Big]ds$$

$$\leqslant \frac{1}{2}C^2 \sup_{0\leqslant k\leqslant t_n} \mathbb{E}[\|D\mathcal{G}(u_{h,k})F(u_{h,k})\|^2_{L^2(D)}]\Delta t^3 \leqslant C_{5,1}\Delta t^3.$$

Finally,

$$\mathbb{E}[\|V_2\|^2_{L^2(D)}] = \frac{1}{4}\mathbb{E}\Big[\|\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{(t_n - t_{k+\frac{1}{2}})A_h}\Big(\int_0^1 (1-s)\eta^\top P_h D^2\mathcal{G}(u_{h,k} + s\eta)\eta ds\Big)dW(s)\|^2_{L^2(D)}\Big].$$

For the last term, firstly we notice that

$$\int_0^1 (1-s)\eta^\top P_h D^2\mathcal{G}(u_{h,k} + s\eta)\eta ds \leqslant \Delta t \max_{0\leqslant s\leqslant 1}\big((1-s)\eta^\top P_h D^2\mathcal{G}(u_{h,k} + s\eta)\eta\big).$$

Moreover, recall that

$$\eta = \big(A_h u_{h,n} + P_h F(u_{h,n})\big)\Delta t + \mathcal{G}(u_{h,n})\Delta W_n.$$

Hence,

$$\mathbb{E}[\|V_2\|^2_{L^2(D)}] \leqslant C_{5,2}(\Delta t^{7/2} + \Delta t^3 + \Delta t^{5/2}).$$

□

Theorem 4.2.9. *Let equation (4.1) have a unique mild solution $u(t)$ under Assumptions 4.0.1, 4.0.4, 4.0.6 and let $u_{h,n+1}$ defined by (4.22) be the numerical approximation of $u(t)$. Then*

$$\Big(\mathbb{E}\big[\big\|u(t) - u_{h,n+1}\big\|^2\big]\Big)^{1/2} \leqslant Ch^2 + Dh^{\zeta} + K\Delta t^{\theta} + L\Delta t^{1/2}|\ln(\Delta t)|,$$

where C, D, K, L are positive constants, independent of Δt, h and $\zeta \in (0, 2]$, $\theta > 1$.

Proof. We start from the following inequality

$$\begin{aligned}\mathbb{E}\big[\big\|u(t) - u_{h,n+1}\big\|^2_{L^2(D)}\big] \leqslant 5\Big(&\mathbb{E}\big[\|I\|^2_{L^2(D)}\big] + \mathbb{E}\big[\|II\|^2_{L^2(D)}\big] \\ &+ \mathbb{E}\big[\|III\|^2_{L^2(D)}\big] + \mathbb{E}\big[\|IV\|^2_{L^2(D)}\big] + \mathbb{E}\big[\|V\|^2_{L^2(D)}\big]\Big).\end{aligned}$$

By the estimates that we have from Lemmas 4.2.5, 4.2.6, 4.2.7, 4.2.8 we get that

$$\begin{aligned}\mathbb{E}\big[\big\|u(t) - u_{h,n+1}\big\|^2_{L^2(D)}\big] \leqslant 5\Big(&Ch^4 + D(\Delta t^{2\theta} + h^{2\zeta}) + K\Delta t^{2\theta} + L\Delta t(\ln(\Delta t))^2 \\ &+ M\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\big[\|u(x,t) - u_{h,k}\|^2_{L^2(D)}\big]ds\Big).\end{aligned}$$

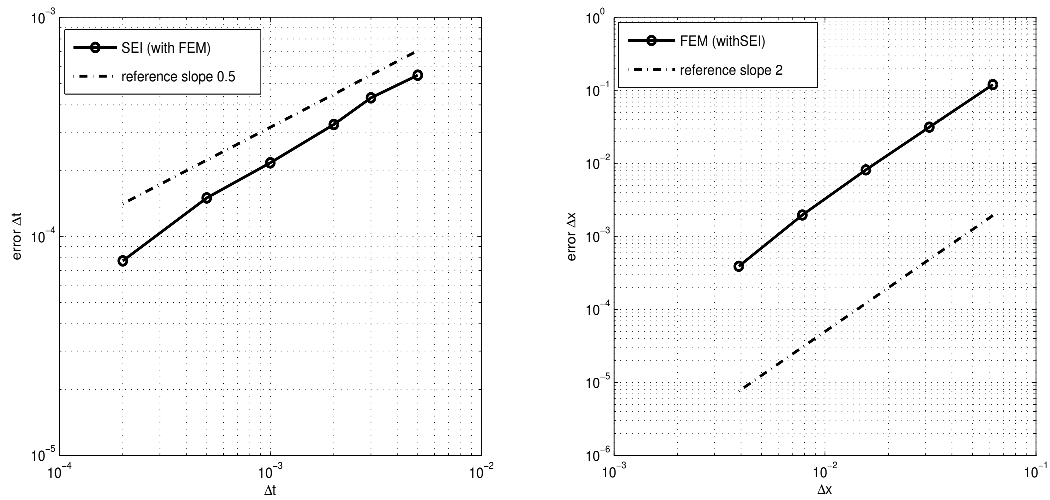
Let $E := Ch^4 + D(\Delta t^{2\theta} + h^{2\zeta}) + K\Delta t^{2\theta} + L\Delta t(\ln(\Delta t))^2$, Gronwall's inequlaity yields

$$\mathbb{E}\big[\big\|u(t) - u_{h,n+1}\big\|^2_{L^2(D)}\big] \leqslant E + ME\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} e^{M\Delta t(t-s)}ds.$$

Finally, taking the square root on both sides of the inequality finishes the proof. □

The convergence result of Theorem 4.2.9 is confirmed by our numerical experiments. We look at the strong convergence of SEI for a Stratonovich-type SPDE with multiplicative noise. Firstly, we fix our mesh in space and we examine convergence for the time discretisation method. We use 10^{-4} as a reference time step and multiples of 10^{-4} as bigger time-steps. We generate M samples and compute the error between the reference solution and the approximate solutions in the mean square sense. As seen in Figure 4.1a, the slope of the line that fits the convergence rate of the approximate solutions to the reference one, has the average value 0.6. This agrees with the $\Delta t^{1/2+\epsilon}$ result which corresponds to the non-diagonal case of our analysis.

Next, we fix our time step and we look at convergence in space for the FEM-method. Our reference solution is taken by 2^9 elements with a reference spatial step $\frac{L}{2^9-1}$ in the domain $D = [0, L]$. Again, we approximate the solution using multiples of the reference spatial step that create coarser meshes and we compute the averaged error between the reference and the approximate solutions for M different samples. The order of convergence, as seen in Figure 4.1b, agrees with Theorem 4.2.9.



(a) Strong convergence of SEI for SPDEs (b) Strong convergence of FEM for SPDEs

Figure 4.1: Order of convergence for SEI and FEM when approximating an SPDE of the form with $f(u) = u(1-u)(u+0.5)$, $g(u) = \sigma u(1-u)$, where $\sigma = 0.5$ is the noise intensity. For (a), reference time step $Dt_{\text{ref}} = 10^{-4}$ with multiples $Dt = \kappa Dt_{\text{ref}}$, $\kappa = (2, 5, 10, 20, 40, 50)$. Fixed spatial step-size 1/63. Slope of the best fitted line, using polyfit, is 0.604180. For (b), reference spatial step $h_{\text{ref}} = 1/(2^9-1)$, with multiples $h = \kappa h_{\text{ref}}$, $\kappa = (2, 2^2, 2^3, 2^4, 2^5)$. Fixed time-step size 10^{-4}. For both Figures $M = 100$ samples.

The example that we use is the stochastic Nagumo equation of the form (4.1), with A being the second order spatial operator u_{xx} and $f(u) = u(1-u)(u+0.5)$, $g(u) = \sigma u(1-u)$. Note that the drift and diffusion terms of the Nagumo model satisfy the Lipschitz assumption only locally. However, we use the model in our numerical experiments for display purposes. Later in Chapter 5, we examine the Nagumo model in more detail and apart from directly simulating its solution, we suggest an alternative approximation method.

Chapter 5

Stochastic travelling waves

In this chapter we are going to see an application of the SEI discretisation scheme in SPDEs. Firstly, we need to consider an individual space and time discretisation. In particular, we start from discretising in space which converts the SPDE into a system of SODEs and then we use SEI for discretising the SDEs in time. At this stage, we recall the general form of the Stratonovich SPDE

$$du = \big(u_{xx} + f(u)\big)dt + g(u) \circ dW,$$

$$u(0, x) = u^0(x). \tag{5.1}$$

Our space and time variables are in finite domains $t \in [0, T]$, $T > 0$, $x \in [0, L]$ and we are going to use Neumann boundary conditions. Then, (5.1) can be rewritten in the form

$$du = \big(\tilde{A}u + f(u)\big)dt + g(u) \circ dW, \tag{5.2}$$

where $\tilde{A}$ is in general an unbounded linear operator that denotes the second spatial derivative ∂_{xx}. Now, if the drift and diffusion terms f and g respectively are globally Lipschitz functions, the mild solution of (5.2), as seen in Chapter 4, is given by

$$u(t) = e^{t\tilde{A}}u_0 + \int_0^t e^{\tilde{A}(t-s)} f(u)ds + \int_0^t e^{\tilde{A}(t-s)} g(u) \circ dW(s). \tag{5.3}$$

We consider W to be a Q-Wiener process and we use the expansion of equation (4.3) to represent it in the following way. If $\{\phi_j\}_{j\in\mathbb{Z}}$ is an orthonormal basis of the Hilbert space $L^2([0, L])$, then

$$W(x, t) = \sum_{j\in\mathbb{Z}} \sqrt{\zeta_j}\phi_j(x)\beta_j(t),$$

where $\beta_j(t)$ are independent, identically distributed (iid) Brownian motions and ζ_j, ϕ_j are the eigenvalues and the corresponding eigenfunctions of the covariance operator Q. Recall from Chapter 4, that there is a connection between the rate of decay of the eigenvalues ζ_j and the smoothness of the noise path in space. In this approach, $\zeta_j = |j|^{-(2s+1+\epsilon)}$ with $s = 1$, $\epsilon > 0$, that is $W(x) \in H^1$ which means that our noise is smooth in space but white in time. Also, we assume that the covariance operator Q and the linear operator ∂_{xx} have the same eigenfunctions ϕ_j.

As mentioned before, for discretising in space we use standard uniformly spaced finite differences on a finite grid $x_0, \ldots, x_m$. So, we have m points taken with spatial step Δx and we approximate the solution at the point x_j, $u(x_j)$ by u_j. Thus, the discretised in space SPDE reads

$$du = \big(A_{\Delta x} u + f(u)\big)dt + g(u) \circ dW \tag{5.4}$$

with $A_{\Delta x}$ as given in (4.7).

For discretising in time, we use the Stratonovich exponential integrators-based scheme (SEI), analysed in Section 2.1. Then, the approximate solution at time $t_n = n\Delta t$, denoted by u_n, is given by modifying (1.43) accordingly

$$u_{n+1} = e^{\Delta t A} u_n + \phi(\Delta t A) f(u_n) + \frac{1}{2} e^{\frac{\Delta t A}{2}} \big(g(u_n) + g(\tilde{u}_{n+1})\big)\Delta W_n,$$

where ΔW_n are the Brownian increments $W(t_{n+1}) - W(t_n)$, ϕ is a function given by $\phi(\Delta t A) = A^{-1}(e^{\Delta t A} - I)$ and $\tilde{u}_{n+1}$ is a predictor which in this approach is taken by another exponential integrators-based scheme,

$$\tilde{u}_{n+1} = e^{\Delta t A} u_n + \phi(\Delta t A) f(u_n) + e^{\Delta t A} g(u_n)\Delta W_n.$$

So, the fully discretised SPDE reads

$$u_{n+1} = e^{\Delta t A_{\Delta x}} u_n + \phi(\Delta t A_{\Delta x}) + \frac{1}{2} e^{\frac{\Delta t A_{\Delta x}}{2}} \big(g(u_n) + g(\tilde{u}_{n+1})\big)\Delta W_n. \tag{5.5}$$

5.1 Simulation of stochastic travelling waves

In this section we use SEI for the simulation of stochastic travelling waves as solutions of the stochastic Nagumo model. The initial deterministic Nagumo model was introduced in [62] and was designed to simulate a nerve axon by means of a reaction-diffusion equation. Depending on its parameter values the system may have travelling wave solutions, see [35, 71], which is the case that we are going to look at. In particular, we consider the stochastic Nagumo equation with Stratonovich multiplicative noise for the case of travelling wave solutions which we look at over truncated domains from $\mathbb{R}$. Looking back at (5.1), in order to get the Nagumo equation, we substitute for $f(u) = u(1-u)(u-0.25)$ and for multiplicative noise, the diffusion term is given by a function that depends on u, here we take $g(u) = u(1-u)$. Note that in the case of $g(u) = 0$, we retrieve the deterministic Nagumo equation which has a travelling wave solution that can be seen in Figure 5.1.

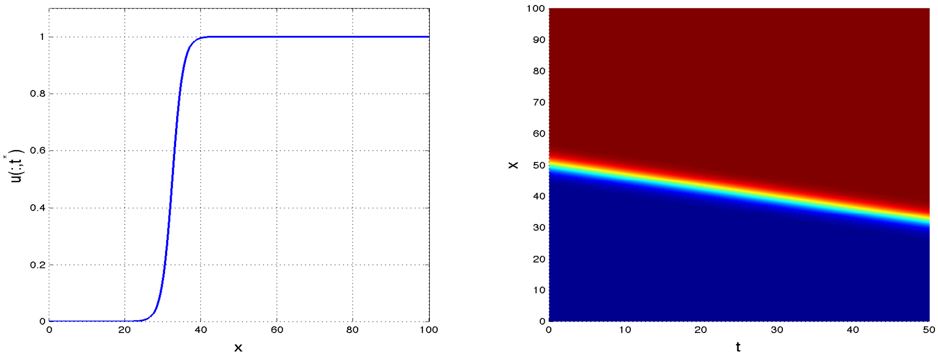


Figure 5.1: Deterministic Nagumo equation, solved for $t \in [0, 50]$, $x \in [0, 100]$ with Neumann BCs and initial condition $u_0(x) = (1 + e^{-x})^{-1}$. Left: wave solution plot against space at fixed final time. Right: time-space plot of the solution showing the wave travelling with constant speed that can be found by computing the slope of the front.

Now, for the stochastic case we consider

$$du = \big(u_{xx} + u(1-u)(u-0.25)\big)dt + u(1-u) \circ dW, \tag{5.6}$$

with $t \in [0, T]$, $x \in [0, L]$ and Neumann boundary conditions. As initial condition we consider the sigmoid function $u_0(x) = (1 + e^{-x})^{-1}$. Then, both $u = 0$ and $u = 1$ make $g(u)$ equal to zero and as indicated by numerical simulations, there exists a stochastic travelling wave between these two values, for a theoretical proof of existence of a stochastic travelling wave solution see [82]. A way to intuitively understand a stochastic travelling wave, is to look at the solution of the (S)PDE against space for some fixed time t^* and think of the deterministic wave driven by some noise, as displayed in Figure 5.2. Note that this is just one realisation of the stochastic wave, each realisation is different since the solution of an SPDE is a random variable. The smoothness of the stochastically forced wave reflects our choice of eigenvalues and corresponding eigenfunctions of the covariance operation Q.

We also need to define the speed of the wave and in order to do this, we introduce two different notions of the speed. The first notion requires looking at the wave solution at a fixed time t_i and also fixing a point of the solution. In this approach we choose to fix the midpoint which is given by 0.5 and is denoted by $m(t)$. Next, we look at the solution at the time t_{i+1}, we measure the displacement of the midpoint and we divide by the step-size. This notion of the speed, denoted by λ_m, is called

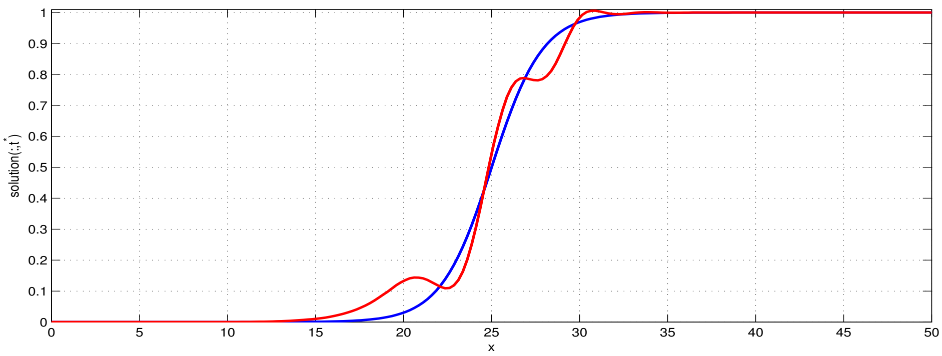


Figure 5.2: Deterministic and stochastic fixed time wave solutions plot against space in blue and red respectively.

instantaneous speed and is given by

$$\lambda_m(t_i) := \frac{m(t_i) - m(t_{i-1})}{\Delta t}, \tag{5.7}$$

where $t_0, \ldots, t_i, \ldots, t_N,\ N \in \mathbb{Z}$ are fixed time steps and Δt is the step-size. As its name suggests, this definition provides the speed of the wave over single time steps. One aspect of the definition that requires special attention is the choice of the fixed point. The reason for this, is that depending on each different realisation of the stochastic wave, some points of the solution might be multivalued, see for example that there are two different values corresponding to the same point 0.1 of the solution u in Figure 5.2.

An alternative way to define the speed of the wave is to consider the following cumulative sum

$$\Lambda_m := \frac{1}{t} \sum_{i=1}^{N} \lambda_m(t_i), \tag{5.8}$$

which gives us the so called time-averaged speed. This could be thought of as an approximation of $\dfrac{d\lambda}{dt}$.

Both (5.7) and (5.8) are defined for a single realisation, but can be easily extended by considering many realisations and then taking the expectation over the number of the realisations

$$\lambda_m^E(t) := \mathbb{E}\big(\lambda_m(t)\big), \qquad \Lambda_m^E := \mathbb{E}\big(\Lambda_m\big).$$

This, gives us the mean instantaneous speed and the mean time-averaged speed over different samples of the solution.

The solution of the stochastic Nagumo equation, given by direct simulation of the SPDE, using FD for discretising in space and SEI for discretising in time can be seen in Figure 5.3. We can observe in Figure 5.3 that unlike the deterministic case, the wave does not travel uniformly in time and also that the speed is not monotonic because of the small oscillations that occur due to the noise at the front. Moreover, the results for the two different notions of the speed as defined in (5.7) and (5.8) can be seen in Figure 5.4.

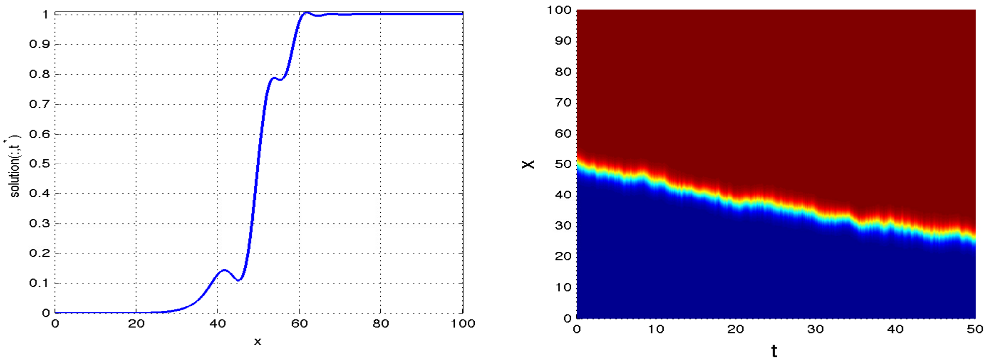


Figure 5.3: Stochastic Nagumo equation, solved for $t \in [0, 50]$, $x \in [0, 100]$ with Neumann BCs and initial condition $u_0(x) = (1 + e^{-x})^{-1}$. Left: stochastic wave solution plot against space for fixed time. Right: time-space plot of the solution with the stochastic wave travelling not uniformly in time.

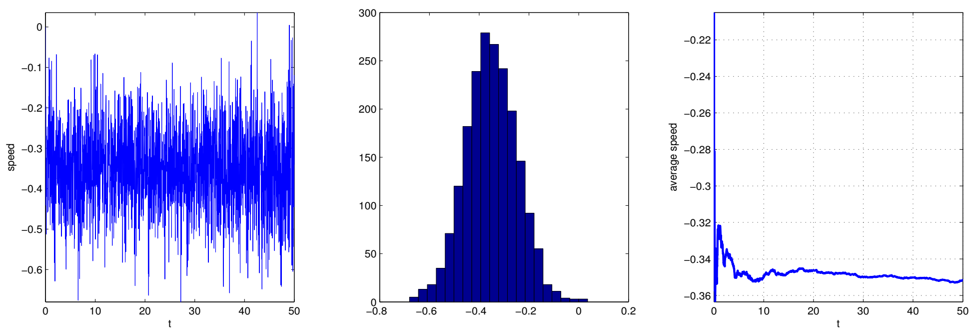


Figure 5.4: From left to right: (a) Instantaneous speed λ_m against time. (b) Distribution of the instantaneous speed λ_m. (c) Time-averaged speed Λ_m against time. All for one realisation of the stochastic travelling wave solution.

Later in Subsection 5.2.1, we consider more realisations of the stochastic travelling wave solution and we plot the mean instantaneous and mean time-averaged speeds too.

5.2 Freezing method for an SPDE

In this section we suggest an alternative method for computing stochastic travelling waves and their speed, which is an extension of [57]. The basic idea of the so called freezing method is that the wave position, and hence the speed, is found by minimising the L^2-distance between a reference function and the travelling wave. Again, we use the stochastic Nagumo equation with Stratonovich multiplicative noise, so that we can compare the results that we get from the two different approaches, but we start from describing the general idea of the method for the deterministic case, as seen in [7].

We consider a PDE of the general form

$$u_t = u_{xx} + f(u), \tag{5.9}$$

in $\mathcal{L}^2(\mathbb{R})$ with $x \in \mathbb{R}$, $t \in \mathbb{R}_+$ and some given initial condition $u(0) = u_0$. We suppose that the PDE has a travelling wave solution $u(x,t)$, also written as $u(\xi)$ where $\xi = x - \lambda t$ and λ denotes the wave-speed. Then, we look at the PDE in a co-moving frame which is obtained by considering the transformation $v(\xi, t) = u(x - \lambda t, t)$, so (5.9) reads

$$v_t = v_{\xi\xi} + \lambda v_\xi + f(v), \qquad \xi \in \mathbb{R}, \quad t > 0, \tag{5.10}$$

note that $u(x,t)$ is a stationary solution for the transformed equation which means that it satisfies (5.10) with 0 on the left hand side. In the case that the speed of the wave is not constant but a function that depends on time, we set the wave position to be a time-dependent function $\gamma(t)$ and the wave-speed is given by the rate $\dfrac{d\gamma(t)}{dt}$, or equivalently $\gamma(t) = \displaystyle\int_0^t \lambda(s)ds$. Then, (5.10) is finally transformed in

$$v_t = v_{xx} + \lambda(t)v_x + f(v). \tag{5.11}$$

Note that (5.11) is still in a co-moving frame and it also contains the unknown wave-speed λ. The idea in order to compensate for λ is to add an extra condition which enables us to couple (5.11) with this added condition and create a system that we solve for the travelling wave solution v and the speed of the wave λ.

In order to derive the extra condition, we align the travelling wave solution to a continuous reference function $\hat{v}$, of fixed shape. Then, we minimise the difference between v and $\hat{v}$ over shifts in space y as suggested in [7], that is we consider

$$\min_y \|v(x,\cdot) - \hat{v}(x-y,\cdot)\|_2^2,$$

which by differentiating with respect to y and equating to zero gives

$$\int_{\mathbb{R}} \big(v(x,\cdot) - \hat{v}(x-y,\cdot)\big)\hat{v}_x dx = 0,$$

which is equivalent to the inner product equation

$$\langle \hat{v}_x, v - \hat{v}\rangle = 0. \tag{5.12}$$

At this stage, we couple (5.11) with the condition (5.12) to get the system

$$v_t = v_{xx} + \lambda(t)v_x + f(v)$$

$$\langle \hat{v}_x, v - \hat{v}\rangle = 0, \tag{5.13}$$

which is a partial differential algebraic equation (PDAE). If we discretise and solve (5.13) for v and λ we get the travelling wave solution and the speed of the wave as seen in Figure 5.5. The fact that the front of the wave now remains now at the same position throughout the computation indicates that the wave is frozen that is it does not travel in time any more.

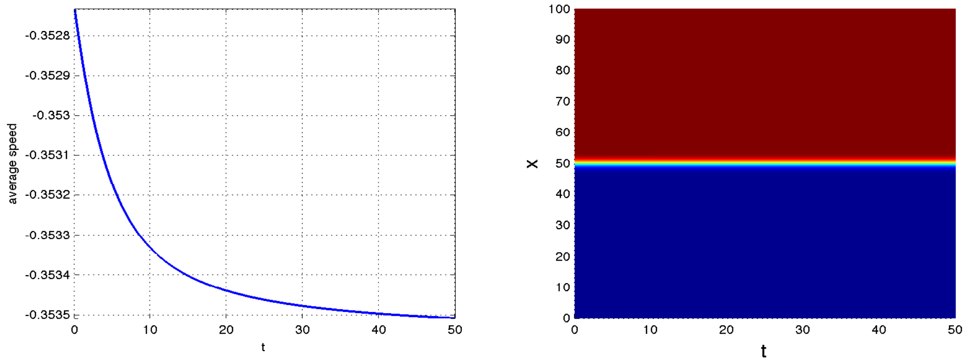


Figure 5.5: Deterministic Nagumo equation, solved for $t \in [0, 50]$, $x \in [0, 100]$ with Neumann BCs and initial condition $u_0(x) = (1 + e^{-x})^{-1}$. Left: wave-speed λ given directly as an output from (5.13). Right: time-space plot of the solution with the frozen wave.

Imitating the deterministic case, we extend the freezing method to the SPDE, as it was done in [57]. We consider the SPDE (5.1) in a co-moving frame with time-dependent wave-speed, by using the transformation $u(x, t) = v(x - \gamma(t), t)$. Then, $du = v_x d\gamma(t) + dv$ and $u_{xx} = v_{xx}$ so (5.1) reads

$$dv = \big(v_{xx} + \frac{d\gamma(t)}{dt}v_x + f(v)\big)dt + g(v) \circ dW,$$

equivalently

$$dv = \big(v_{xx} + \lambda(t)v_x + f(v)\big)dt + g(v) \circ dW. \tag{5.14}$$

We couple (5.14) with (5.12) to convert the SPDE into a stochastic partial differential algebraic equation (SPDAE)

$$dv = \big(v_{xx} + \lambda(t)v_x + f(v)\big)dt + g(v) \circ dW$$

$$0 = \langle \hat{v}_x, v - \hat{v}\rangle \tag{5.15}$$

For discretising in space we use the FD method and for discretising in time we use the SEI scheme, so our fully discretised scheme is given by

$$v_{n+1} = e^{\Delta t A_{\Delta x}} v_n + \phi(\Delta t A_{\Delta x})\Big(\lambda_{n+1} D_c v_n + f(v_n)\Big) + \frac{1}{2} e^{\frac{\Delta t A_{\Delta x}}{2}} \Big(g(v_n) + g(\tilde{v}_{n+1})\Big)\Delta W_n$$

$$0 = (D_C \hat{v})^\top v_{n+1} - (D_C \hat{v})^\top \hat{v}, \tag{5.16}$$

where $D_C \approx \partial_x$. Note that for the numerical implementation we write the discretised scheme in the following matrix form

$$\begin{pmatrix} I & -\phi(\Delta t A) D_c v_n \\ (D_c \hat{v})^\top & 0 \end{pmatrix} \begin{pmatrix} v_{n+1} \\ \lambda_{n+1} \end{pmatrix} =$$

$$\begin{pmatrix} e^{\Delta t A_{\Delta x}} v_n + \phi(\Delta t A_{\Delta x}) f(v_n) + \frac{1}{2} e^{\frac{\Delta t A_{\Delta x}}{2}} \Big(g(v_n) + g(\tilde{v}_{n+1})\Big)\Delta W_n \\ (D_c \hat{v})^\top \hat{v} \end{pmatrix}.$$

For further reading related to the theory and simulation of deterministic DAEs we refer to [60], while for SDAEs we refer to [73, 84] where there is an application of SDAEs to the transient noise simulation of electronic circuits. Finally, we refer to [49] where a Runge-Kutta method is used for approximating solutions to SDAEs.

5.2.1 The stochastic Nagumo model

As a first example to illustrate how the freezing method works, here we solve the stochastic Nagumo equation with Stratonovich multiplicative noise. Then, we compare the results that we get by the two different approaches, direct simulation of the SPDE and the freezing method. As we see in Figure 5.6 left, in the SPDE case the front of the wave starts at a position $x = 50$ and it moves down with the passage of time. By contrast, in the SPDAE approach the wave starts at the same position and it remains on average at this position throughout the computation, which as in the deterministic case indicates that the stochastic wave is frozen.

We can also compare the results that we have for the two different notions of the speed, given by the two different approaches, as they can be seen in Table 5.1. Recall that we denote the travelling instantaneous and time-averaged speeds by λ_m and Λ_m respectively, whereas frozen instantaneous and time-averaged speeds are denoted by λ and Λ respectively. Both the instantaneous speed and the time-averaged speed values are taken from one realisation of the stochastic wave at the final computational time-step. Note that the analytic wave speed for the deterministic equation is given by $\Lambda = -0.353553$.

Table 5.1: Instantaneous and time-averaged speeds of both the travelling and the frozen wave cases, for a single realisation of the solution.

Speed	Stochastic travelling wave	Stochastic frozen wave
Instantaneous speed	λ_m=-0.340524	λ=-0.327001
Time-averaged speed	Λ_m=-0.351980	Λ=-0.348821

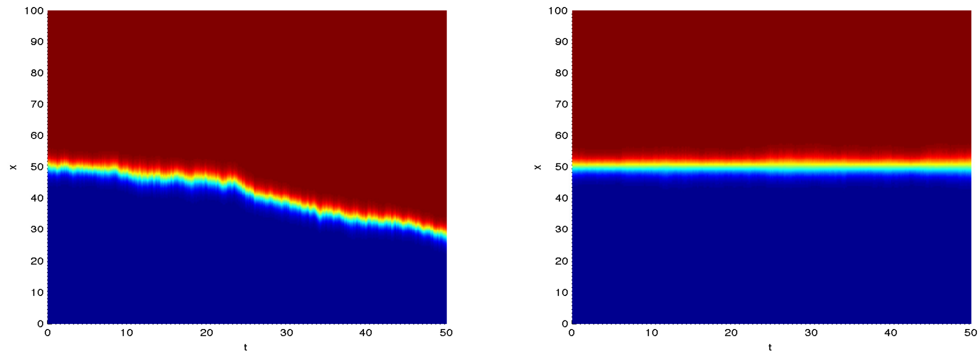


Figure 5.6: Stochastic Nagumo equation, solved for $t \in [0, 50]$, $x \in [0, 100]$ with Neumann BCs and initial condition $u_0(x) = (1 + e^{-x})^{-1}$. Left: time-space plot of the travelling wave given by direct simulation of the SPDE. Right: time-space plot of the solution with the frozen wave given by the freezing method which solves the SPDAE.

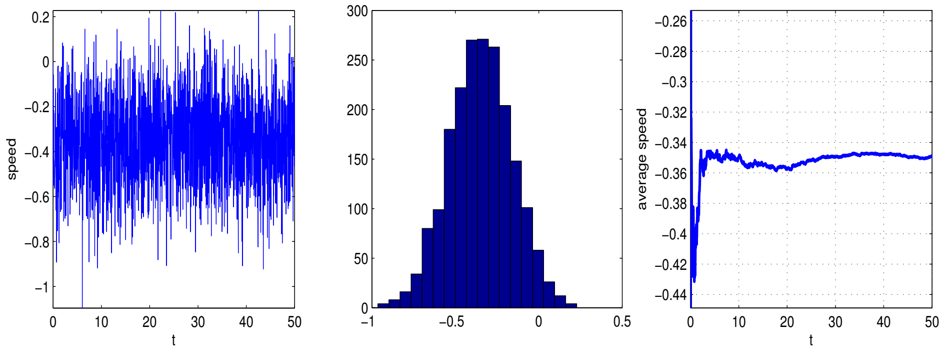


Figure 5.7: Given as an output from solving the SPDAE, from left to right: (a) Instantaneous speed λ against time. (b) Distribution of the instantaneous speed λ. (c) Time-averaged speed Λ against time, all for one realisation of the frozen stochastic wave.

Another way to compare the values of the time-averaged speed in the case of the SPDE and the SPDAE is to plot them together against time, Figure 5.8. Next, we can then extend our speed comparisons by considering many realisations of the stochastic wave solution and then taking the expectation over the number of these realisations. This, gives us a clearer view of how close the results, taken by the two different approaches, are even when the noise intensities are increased. After considering many realisations we can also plot the distance of the standard deviation above and below the mean values, Figure 5.9.

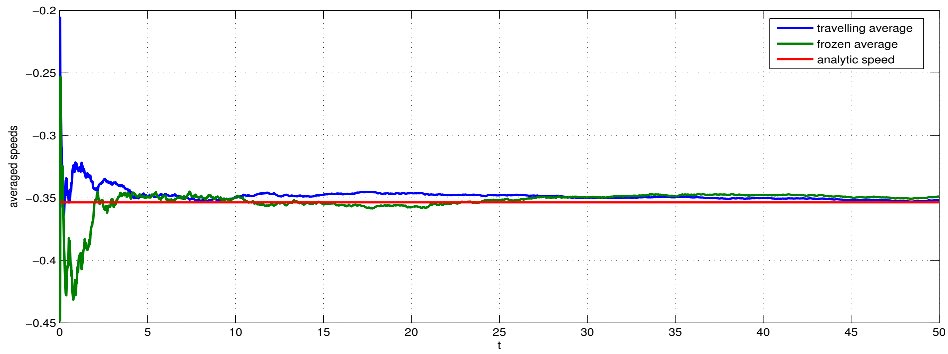


Figure 5.8: Time-averaged speed of the stochastic travelling and frozen wave Λ_m, Λ in blue and green respectively. Noise intensity is 0.1 for the stochastic waves, while the analytic wave speed for the deterministic equation is given in red.

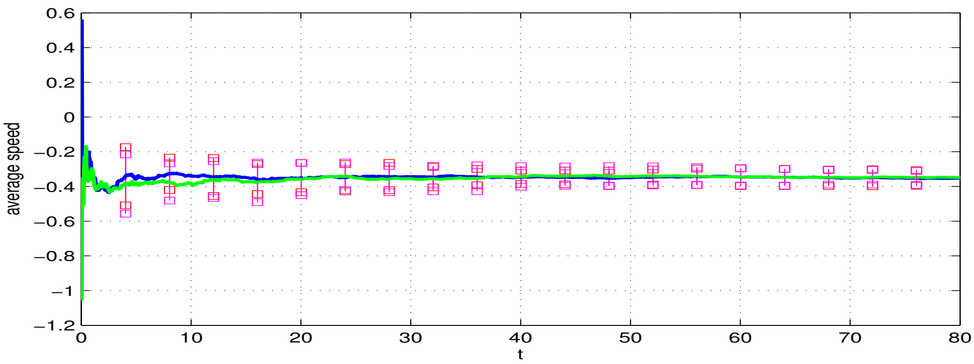


Figure 5.9: Mean time-averaged speed with error bars. Λ^E for the frozen wave in blue and Λ^E_m for the travelling in green. Mean taken over 100 samples with noise intensity 0.5.

To sum up, in this section we introduced the freezing method which requires coupling the SPDE with an extra minimisation condition and thus converting it into a SPDAE. We saw that the results that we get from solving the SPDE and the SPDAE are comparable, which means that we can use the freezing method for computing stochastic travelling waves and in that case we do not need a large computational domain for capturing the behaviour of the wave. Moreover, we have the wave-speed given as an output from solving the algebraic equation and we can also get some statistics regarding the wave-speed. The cost for adding the extra condition is just the computation of an inner product, which is what the extra condition essentially is.

Another remark that is worth mentioning is that as we increase the noise intensity in our system, the magnitude of the wave-speed increases as well, Figure 5.10. The opposite behaviour is observed in the Itô case, as seen in [57], where the speed-magnitude decreases as we increase the noise intensity.

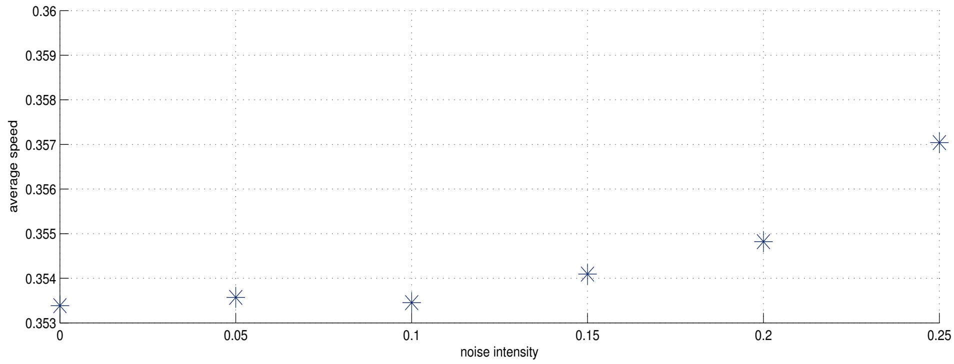


Figure 5.10: Time-averaged speed at the final computational time against noise intensity. Mean values taken over 1000 different realisations. Speed magnitude increasing as the noise intensity increases.

5.2.2 The stochastic FitzHugh-Nagumo model

We can now apply the freezing technique to the stochastic FitzHugh-Nagumo (FHN) model. The FHN model is a simplified two-variable version of the Hodgkin-Huxley model which is often used in neuroscience to describe the initiation and the propagation of a neuronal signal, that is of an electrical pulse often also called action potential. The deterministic FHN model is given by the following system

$$\begin{aligned} du &= \big(u_{xx} + u(1-u)(u-\alpha) - z\big)dt \qquad (5.17) \\ dz &= \epsilon(u - \gamma z)dt, \end{aligned}$$

where u is the so called fast variable associated with the potential, α, γ, ϵ are constants with $0 < \alpha < 1$ and $\epsilon \ll 1$. A wave solution of the system (5.17) has the form $\big(u(x,t), z(x,t)\big)$ that can also be written as $\big(u(\xi), z(\psi)\big)$ with $\xi = x - \lambda t$ and $\psi = x - \lambda t$ where λ denotes the speed of the waves of the components u and z respectively.

For different approaches on various noise FHN models we refer to [6] and references therein. Here, we consider the FitzHugh-Nagumo equation with Stratonovich multiplicative noise as given by

$$du = \big(u_{xx} + u(1-u)(u-0.25) - z\big)dt + \sigma u(1-u) \circ dW \qquad (5.18)$$

$$dz = \big(\epsilon u - \epsilon\gamma z\big)dt,$$

in finite space and time intervals, $x \in [0, L]$ and $t \in [0, T]$, with initial condition $(u_0(x), z_0(x))$ and with Neumann boundary conditions. The diffusion term is given by the function $g(u) = \sigma u(1-u)$, where σ is a parameter that regulates the noise intensity. The travelling wave solution of (5.18) can be seen in Figure 5.11, where we plot u and z against space at four different fixed times. It is then easy to see the displacement of the pulse of the wave. A time-space plot of the travelling wave solutions both for the variable u and z can be seen in Figure 5.12. In Figure 5.13, we plot the time averaged speed of the travelling wave for one realisation. Later, see Figure 5.14, we consider more realisations in order to compare the results obtained by direct simulation with the results obtained by the alternative approach that follows.

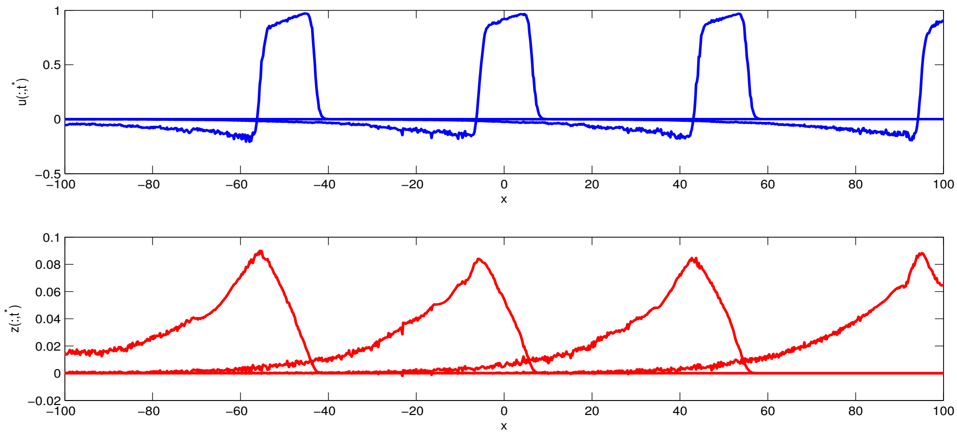


Figure 5.11: Travelling wave solution variables u (top) and z (bottom) against space, at fixed times $t^* = 0, T/4, T/2, 3T/4, T$ for $t \in [0, T]$. We solve FHN system (5.18) with noise intensity $\sigma = 0.05$.

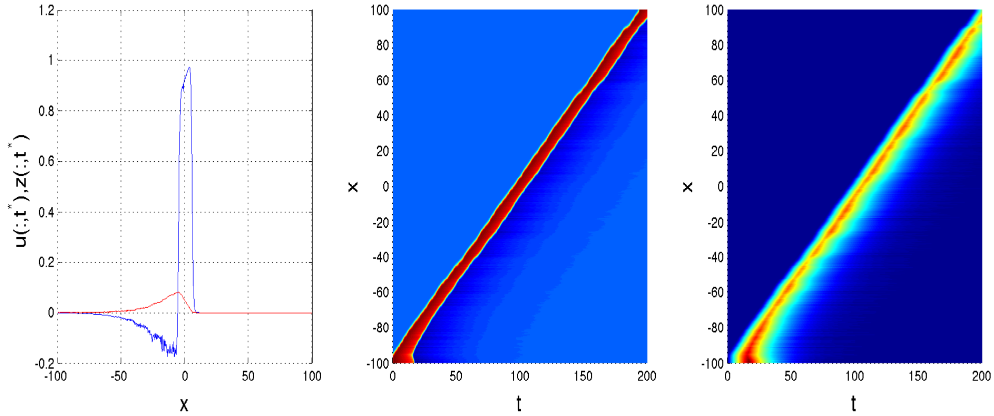


Figure 5.12: Approximate solution of the system (5.18) with noise intensity 0.05. From left to right, solutions u and z against space at fixed time $t^* = T/2$. Time-space plots of the travelling wave solutions u and z.

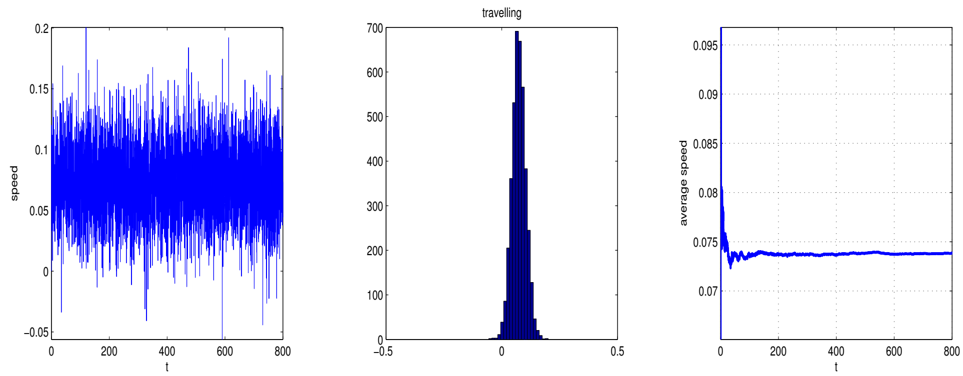


Figure 5.13: From left to right: (a) Instantaneous speed λ_m against time. (b) Distribution of the instantaneous speed λ_m. (c) Time-averaged speed Λ_m against time. All for one realisation of the stochastic travelling wave solution.

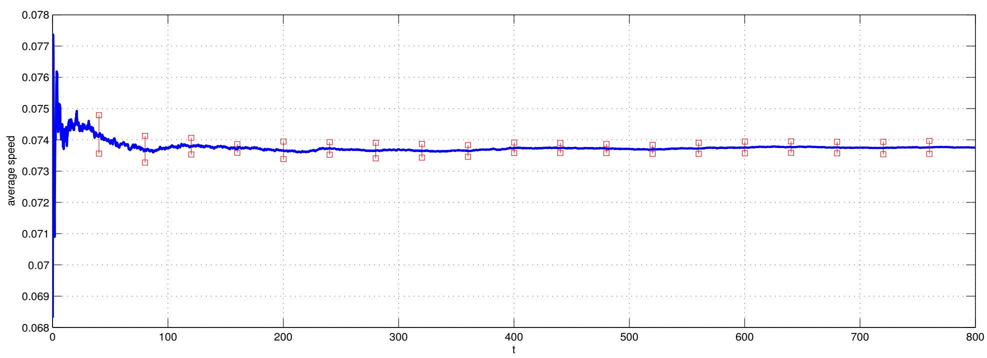


Figure 5.14: Mean of the time-averaged speed $\mathbb{E}(\Lambda_m)$ of the travelling wave plot against time. The mean, taken over 20 samples is given by 0.073279.

In a similar way to the Nagumo equation we look at (5.18) in a co-moving frame by considering the transformation $v(\xi, t) = u(x - \lambda t, t)$ and $w(\psi, t) = z(x - \lambda t, t)$.

Then (5.18) reads

$$dv = \big(v_{xx} + \lambda(t)v_x + v(1-v)(v-0.25-w)\big)dt + g(v)\circ dW \tag{5.19}$$

$$dw = \big(\epsilon v - \epsilon\gamma w + \lambda(t)w_x\big)dt.$$

Now the system (5.19) contains the unknown λ as well. We compensate for $\lambda(t)$ by adding an algebraic constraint, derived in the same way with (5.12) for some continuous reference function $\hat{v}$. Hence, the system (5.19), transformed into an SPDAE in compact vector form, reads

$$d\mathbf{u} = \big(A\mathbf{u} + \lambda(t)\mathbf{u}_x + \mathbf{F}(\mathbf{u})\big)dt + G(\mathbf{u})\circ dW \tag{5.20}$$

$$0 = \langle \hat{\mathbf{u}}_x, \mathbf{u} - \hat{\mathbf{u}}\rangle,$$

where

$$\mathbf{u} = \begin{pmatrix} v \\ w \end{pmatrix}, \quad \mathbf{F}(\mathbf{u}) = \begin{pmatrix} v(1-v)(v-0.25)-w \\ \epsilon v - \epsilon\gamma w \end{pmatrix}, \quad G(\mathbf{u}) = \begin{pmatrix} \sigma v(1-v) \\ 0 \end{pmatrix},$$

and $\mathbf{u}_x, A$ denote the first and second order spatial differential operators respectively. After discretising with FD in space and SEI in time, we obtain

$$\mathbf{u}_{n+1} = e^{\Delta t A_{\Delta x}}\mathbf{u}_n + \phi(\Delta t A_{\Delta x})\big(\lambda_{n+1}D_c\mathbf{u}_n + \mathbf{F}(\mathbf{u}_n)\big) + \frac{1}{2}e^{\frac{\Delta t A_{\Delta x}}{2}}\big(G(\mathbf{u}_n) + G(\tilde{\mathbf{u}}_{n+1})\big)\Delta W_n \tag{5.21}$$

$$0 = (D_c\hat{\mathbf{u}})^\top \mathbf{u}_{n+1} - (D_c\hat{\mathbf{u}})^\top \hat{\mathbf{u}}.$$

In a similar way with (5.16), we can write (5.21) in a matrix form and solve it for the frozen wave solution and the speed of the waves λ.

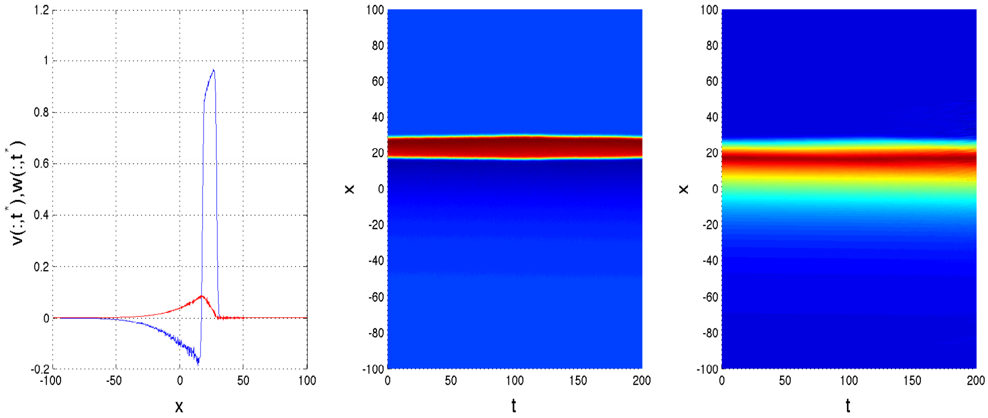


Figure 5.15: Freezing method used for approximating FHN model (5.19) with noise intensity $\sigma = 0.1$. From left to right, solutions v and w against space at the final approximating time. Time-space plots of the frozen wave solutions v and w.

We solve the system (5.20), to obtain the solutions v and w together with the corresponding speed of the waves, λ. The fixed-wave solution space-time plots can be seen in Figure 5.15, while in Figure 5.16 we see the solutions plot against space for different fixed times. The speed of the wave, given as an output from solving the SPDAE, can be seen in Figure 5.20, while in Figure 5.21, we plot the time averaged frozen speed with error bars. In contrast to Figure 5.12 where the waves are travelling in time, we observe here that the pulse of the wave starts at a certain position x and remains on average at the same position for both the solutions. Another thing that confirms that the wave is frozen, is that the midpoint definition of equation (5.8) when applied to the frozen wave, gives a zero-speed, as seen in Figure 5.19.

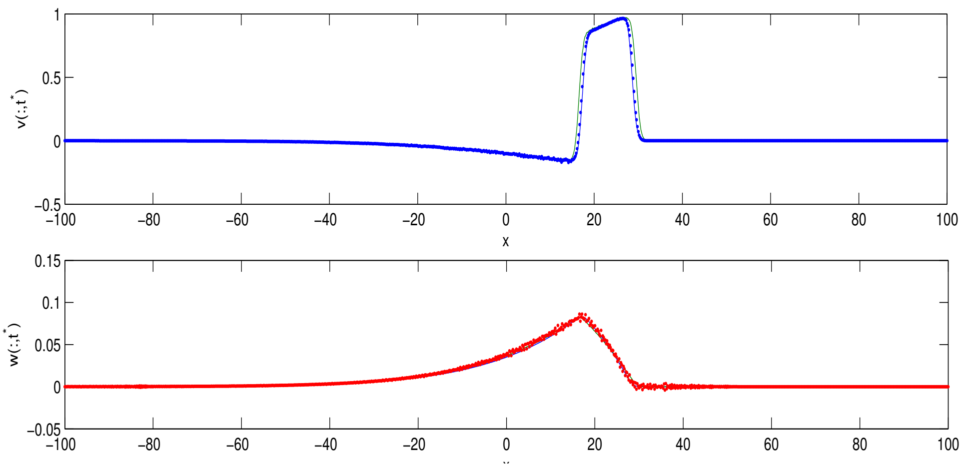


Figure 5.16: From top to bottom, plot of the solutions v and w of the FHN model against space at fixed times $t^* = 0, T/2, T$ for $t \in [0, T]$. Noise intensity $\sigma = 0.1$.

Finally, we approximate the solution of (5.18) for a longer time, by using direct simulation, as seen in Figure 5.17 and my means of the freezing method, as seen in Figure 5.18. Comparing the two different figures, we observe that in the direct simulation case, we fail to capture the behaviour of the wave, as it travels and it gets out of our computational domain. When using the freezing method on the other hand, the pulse remains on average at the same position, enabling the wave to stay within our computational domain. This confirms the claim that the freezing method performs more efficiently when considering long time simulations.

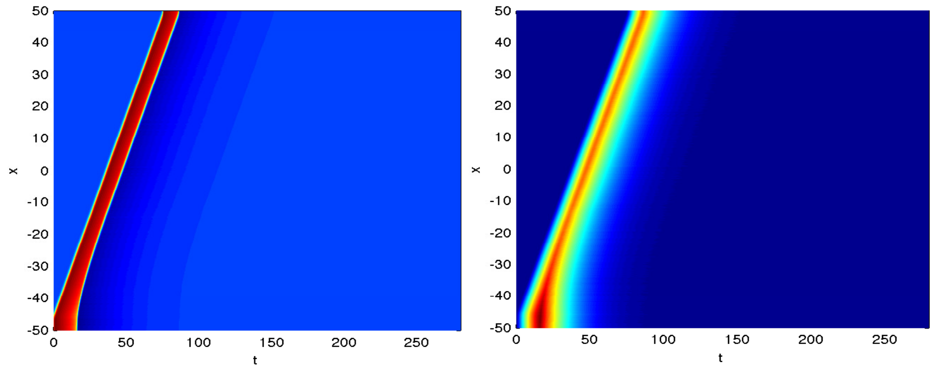


Figure 5.17: Travelling solution for FHN model (5.18) with noise intensity $\sigma = 0.01$. From left to right, solutions u and z against space at the final approximating time. Time-space plots of the travelling wave solutions u and z.

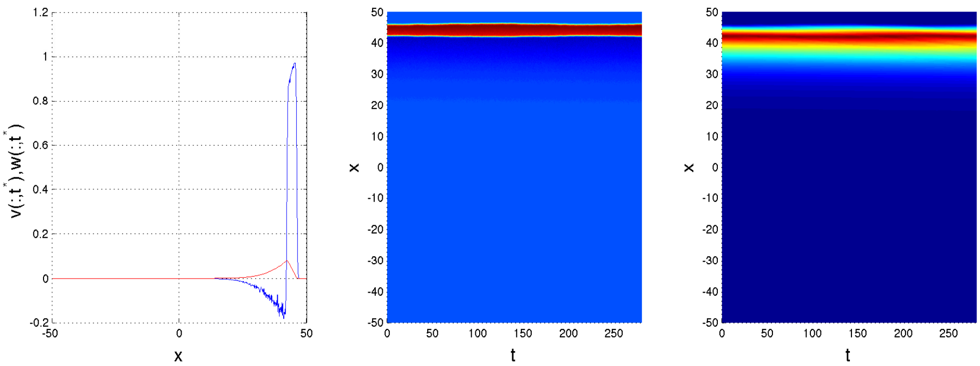


Figure 5.18: Frozen solution for FHN model (5.19) with noise intensity $\sigma = 0.01$. From left to right, solutions v and w against space at the final approximating time. Time-space plots of the frozen wave solutions v and w.

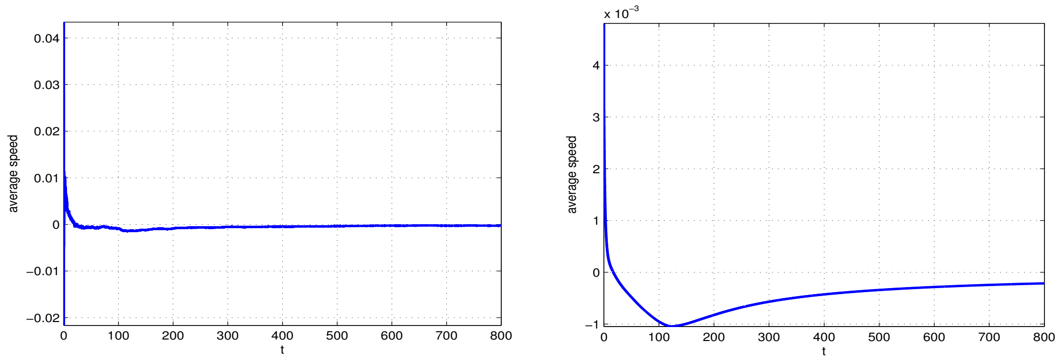


Figure 5.19: Time-averaged speed Λ of the stochastic frozen wave on the left and the deterministic frozen wave on the right. Λ is computed here using the midpoint definition of equation (5.8).

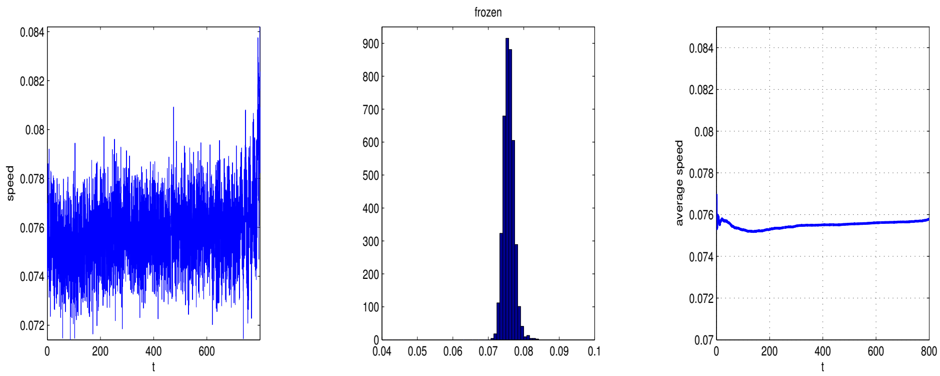


Figure 5.20: Given as an output of the SPDAE, from left to right for a single realisation: (a) Instantaneous speed of the solution variable $\mathbf{u}$ against time. (b) Distribution of the instantaneous speed of $\mathbf{u}$. (c) Time-averaged speed of $\mathbf{u}$ against time.

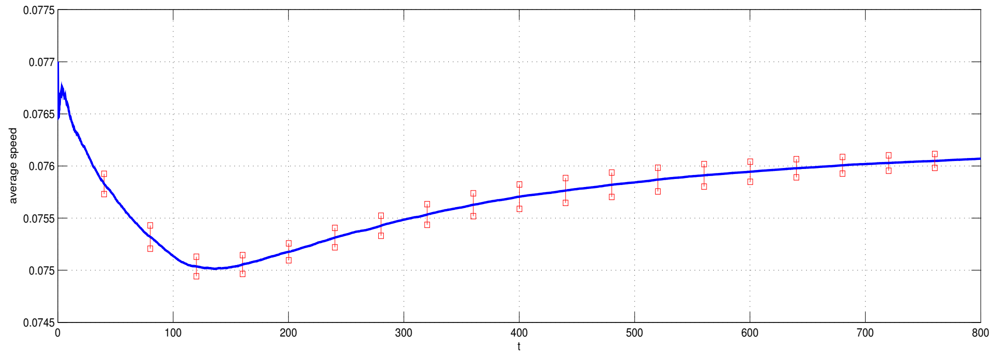


Figure 5.21: Mean of the time-averaged speeds of the frozen wave. The mean, taken over 20 samples, is given by 0.75356.

In a similar way with the Nagumo equation, we plot the instantaneous and time-averaged speed together with a histogram of the instantaneous speed as a random variable, as seen in Figure 5.20. Finally, we plot the mean of the time-averaged speed $\mathbb{E}(\Lambda)$ of the frozen wave, taken over a number of different samples, see Figures 5.14 and 5.21.

We conclude that the results obtained by the two different approaches, i.e. direct simulation and freezing method are at least comparable. More particularly, the mean time-averaged speeds differ for about 0.002, as seen in Figures 5.14 and 5.21. Note that this difference exists even for the deterministic case where we consider no noise at all. This suggests that the freezing method may be subject to bias due to the computation of the term $\mathbf{u}_x$. Hence, some further investigation is required e.g. by considering an upwind or flux limiter scheme.

Chapter 6

Conclusion and future work

The aim of this thesis has mainly been to investigate the numerical approximation of Stratonovich interpreted SDEs and SPDEs.

In Chapter 1, we provided the necessary background for understanding what a stochastically forced differential equation is. More specifically, we established the mathematical framework within which existence and uniqueness of solutions to Itô and Stratonovich SDEs exist. Moreover, we reviewed several numerical schemes that are used for approximating solutions to either Itô or Stratonovich SDEs and we gave the definitions of two different notions of convergence.

In Chapter 2, we proved analytically that the SEI scheme converges with a strong order of $O(\Delta t^{1/2+\epsilon})$ $\epsilon > 0$, for the case of d-dimensional semilinear Stratonovich SDEs with general i.e. non-diagonal and not commutative, multiplicative noise. Our numerical experiments, where we considered the approximation of the stochastic LLG equation, confirm the analytical result. However, it appears that the SEI scheme fails to preserve the norm condition that needs to be satisfied for the LLG equation, which leaves some room for experimenting with introducing a projection in order to preserve the norm.

In Chapter 2, we developed the higher order scheme MSEI suited for Stratonovich SDEs and we proved analytically that the scheme converges with a strong order ≈ 1, for the case of general multiplicative noise. Being derived based on exponential integrator techniques, implies that the scheme is efficient for solving stiff problems, while being of Milstein type ensures the higher order of convergence. An approach to be considered as possible future work, would be the use of ETD instead of the Euler-Maruyama scheme at the derivation of MSEI.

Chapter 3 provided a generalisation of the SEI scheme, based on the generalisation of the notion of the stochastic integral. The new scheme is suited for SDEs interpreted according to Itô, Stratonovich and the in-between calculi. Furthermore, we developed a generalised expression for the predictor used in the scheme and we proved analytically that the order of strong convergence for the generalised version of SEI is affected neither by the predictor nor by the calculus choice and remains $O(\Delta t^{1/2} + \epsilon)$.

We compared the generalised version of SEI with the (θ, σ)-Milstein scheme from the literature for approximating solutions to the Heston model. We observed that $(1, 1)$-Milstein approximation agrees with the SEI approximation for $\alpha = 0.5$ which corresponds to the Stratonovich interpretation. We notice however that the terms in the Heston model satisfy the required for our analysis globally Lipschitz and growth assumptions only in truncated intervals locally. The use of tamed methods has been suggested as a solution to overcome this obstacle. Adapting the SEI scheme accordingly so that the globally Lipschitz assumption can be reduced to a local assumption, seems to be a natural possible direction to consider.

In Chapter 4, we extended the use of SEI scheme to the SPDEs framework. In particular, we considered second order semilinear reaction diffusion Stratonovich SPDEs with finite dimensional multiplicative trace-class noise which is white in time but has some smoothness in space. In Theorem 4.2.9 we proved strong convergence when using SEI and FEM as temporal and spatial discretisation respectively. Our numerical strong convergence results are in agreement with the theoretical results of Theorem 4.2.9. As a possible future direction we could consider modifying the scheme by splitting the linear and nonlinear reaction term. It would then be interesting to investigate convergence and stability properties under the modification.

In Chapter 5, we used the SEI scheme combined with finite differences for simulating stochastic travelling wave solutions for the FitzHugh-Nagumo model. Also, we defined two different notions of the speed which we used for computing the speed of the waves. Alternatively, we suggested the use of the freezing method for approximating the solutions of the stochastic FHN model. The freezing method consisted of minimising the L^2-distance between a chosen reference function and the travelling wave solution, in other words coupling the SPDE with an algebraic constraint and hence converting it into an SPDAE. We found that the numerical results obtained by direct simulation of the models and by means of the freezing method agree but by solving the SPDAE we benefit both from fixing the wave and from getting the speed of the wave as an output from the system without having to compute it separately.

One of the things that would be interesting to look at, is the convergence of the SEI scheme for the SPDAE case. Although we do not have an analytical proof, our numerical experiments indicate that the order of convergence would not be affected by coupling the SPDE with the constraint. From a numerical perspective, it would be interesting to consider adapting the freezing method algorithm by changing the dynamics of the wave speed.

Chapter 7

Appendix

This chapter mainly consists of some results and some more technical steps that support in some way either the analysis or the simulations of the previous chapters. Section 7.1 provides a short introduction to tensors and the related notation that we use throughout the thesis. Moreover, we give some insight into the way that we perform operations with tensors, arranged in an examples-format. Section 7.2 contains some proofs of lemmas that we present in Chapters 1 and 3. Finally, in Section 6.3, we give some Matlab codes that we use for simulations throughout the thesis.

7.1 Tensors

We start by defining the first and second derivatives of a function $G : \mathbb{R} \to \mathbb{R}^d$ of the form

$$G(u) = \begin{pmatrix} g_1(u) \\ \vdots \\ g_d(u) \end{pmatrix},$$

where $g_i(u) : \mathbb{R} \to \mathbb{R}$, $i = 1, \dots, d$. The first and second derivatives, denoted by ∇G and $\nabla^2 G$ or by DG and $D^2 G$ respectively, are given by the following equations. Although in this background-reading approach we use the '∇' notation, we use DG and $D^2 G$ in our analysis.

$$\nabla G(u) = \begin{pmatrix} \frac{\partial g_1(u)}{\partial u} \\ \vdots \\ \frac{\partial g_d(u)}{\partial u} \end{pmatrix}, \ \nabla^2 G(u) = \begin{pmatrix} \frac{\partial^2 g_1(u)}{\partial u^2} \\ \vdots \\ \frac{\partial^2 g_d(u)}{\partial u^2} \end{pmatrix}.$$

Next, we define the first and second derivatives of a vector function $\mathbf{G} : \mathbb{R}^d \to \mathbb{R}^d$ of the form

$$\mathbf{G}(\mathbf{u}) = \begin{pmatrix} g_1(\mathbf{u}) \\ \vdots \\ g_d(\mathbf{u}) \end{pmatrix},$$

where $g_i(\mathbf{u}) : \mathbb{R}^d \to \mathbb{R}$. The derivatives are given by

$$\nabla \mathbf{G}(\mathbf{u}) = \begin{pmatrix} \frac{\partial g_1(\mathbf{u})}{\partial u_1} & \dots & \frac{\partial g_1(\mathbf{u})}{\partial u_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_d(\mathbf{u})}{\partial u_1} & \dots & \frac{\partial g_d(\mathbf{u})}{\partial u_d} \end{pmatrix} = \begin{pmatrix} \nabla g_1(\mathbf{u}) \\ \vdots \\ \nabla g_d(\mathbf{u}) \end{pmatrix} \in \mathbb{R}^{d\times d}$$

and

$$\nabla^2 \mathbf{G}(\mathbf{u}) = \begin{pmatrix} \nabla\frac{\partial g_1(\mathbf{u})}{\partial u_1} & \dots & \nabla\frac{\partial g_1(\mathbf{u})}{\partial u_d} \\ \vdots & \ddots & \vdots \\ \nabla\frac{\partial g_d(\mathbf{u})}{\partial u_1} & \dots & \nabla\frac{\partial g_d(\mathbf{u})}{\partial u_d} \end{pmatrix} = \begin{pmatrix} \nabla^2 g_1(\mathbf{u}) \\ \vdots \\ \nabla^2 g_d(\mathbf{u}) \end{pmatrix} \in \mathbb{R}^{d^3},$$

where

$$\nabla\frac{\partial g_i(\mathbf{u})}{\partial u_j} = \begin{pmatrix} \frac{\partial^2 g_i}{\partial u_j \partial u_1} \\ \vdots \\ \frac{\partial^2 g_i}{\partial u_j^2} \\ \vdots \\ \frac{\partial^2 g_i}{\partial u_j \partial u_d} \end{pmatrix}.$$

Finally, we define the derivatives of a matrix-function $G : \mathbb{R}^d \to \mathbb{R}^{d\times m}$, with $g_{ij} : \mathbb{R}^d \to \mathbb{R}$, $i = 1, \dots, d$, $j = 1, \dots, m$ of the form

$$G(\mathbf{u}) = \begin{pmatrix} g_{11}(\mathbf{u}) & \dots & g_{1m}(\mathbf{u}) \\ \vdots & \ddots & \vdots \\ g_{d1}(\mathbf{u}) & \dots & g_{dm}(\mathbf{u}) \end{pmatrix}$$

to be

$$\nabla G(\mathbf{u}) = \begin{pmatrix} \nabla g_{11}(\mathbf{u}) & \dots & \nabla g_{1m}(\mathbf{u}) \\ \vdots & \ddots & \vdots \\ \nabla g_{d1}(\mathbf{u}) & \dots & \nabla g_{dm}(\mathbf{u}) \end{pmatrix} \in \mathbb{R}^{d^3}$$

and

$$\nabla^2 G(\mathbf{u}) = \begin{pmatrix} \nabla^2 g_{11}(\mathbf{u}) & \dots & \nabla^2 g_{1m}(\mathbf{u}) \\ \vdots & \ddots & \vdots \\ \nabla^2 g_{d1}(\mathbf{u}) & \dots & \nabla^2 g_{dm}(\mathbf{u}) \end{pmatrix} \in \mathbb{R}^{d^4}.$$

We say that $\nabla G(\mathbf{u})$ or $DG(\mathbf{u}) : \mathbb{R}^{d\times d} \to \mathbb{R}^{d^3}$ with the alternative notation, is a tensor of rank three. A way to think of $DG(\mathbf{u})$ is as a matrix with vectors in its entries. Similarly for $\nabla^2 G(\mathbf{u})$ or $D^2 G(\mathbf{u})$, we say that it is in $\mathbb{R}^{d^4}$ and is a tensor of rank four that can be considered as a matrix with $d \times d$ matrices as entries. Alternatively, we could say that $DG(\mathbf{u}_n)$ and $D^2G(\mathbf{u}_n)$ are the corresponding Jacobian and Hessian matrices in higher dimensions.

We multiply the $DG(\mathbf{u})$ matrix by some vector in the following way

$$DG(\mathbf{u}) \cdot \mathbf{b} = \begin{pmatrix} \nabla g_{11}^\top \mathbf{b} & \dots & \nabla g_{1d}^\top \mathbf{b} \\ \vdots & \ddots & \vdots \\ \nabla g_{d1}^\top \mathbf{b} & \dots & \nabla g_{dd}^\top \mathbf{b} \end{pmatrix},$$

for some $\mathbf{b} \in \mathbb{R}^d$. Note that $DG(\mathbf{u}) \cdot \mathbf{b}$ is not in $\mathbb{R}^{d^3}$, it is in $\mathbb{R}^{d\times d}$ and its entries are scalars that depend on $\mathbf{u}$. More details on how to perform operations with tensors can be seen in the following examples.

Example 3. *In this example we expand the double integral of equation (1.16) in order to obtain a component expression.*

Let $I_i(r,s)$ denote the i-th element of the integral $\int_r^s d\,\boldsymbol{W}(\tau)$ and $I_i(r,t)$ denote the i-th element of the integral $\int_r^t d\,\boldsymbol{W}(s)$, $i = 1,\dots,d$. We start from expanding the inside integral $G(\boldsymbol{u}(r)) \int_r^s d\,\boldsymbol{W}(\tau)$ in the following way

$$\begin{pmatrix} G_{11}(\boldsymbol{u}) & \dots & G_{1d}(\boldsymbol{u}) \\ \vdots & \ddots & \vdots \\ G_{d1}(\boldsymbol{u}) & \dots & G_{dd}(\boldsymbol{u}) \end{pmatrix} \begin{pmatrix} I_1(r,s) \\ \vdots \\ I_d(r,s) \end{pmatrix} = \begin{pmatrix} G_{11}(\boldsymbol{u})I_1(r,s) + \ldots + G_{1d}(\boldsymbol{u})I_d(r,s) \\ \vdots \\ G_{d1}(\boldsymbol{u})I_1(r,s) + \ldots + G_{dd}(\boldsymbol{u})I_d(r,s) \end{pmatrix}.$$

We can now substitute for $G(\boldsymbol{u}(r)) \int_r^s d\,\boldsymbol{W}(\tau)$ so that $DG(\boldsymbol{u}(r))\Big(G(\boldsymbol{u}(r)) \int_r^s d\,\boldsymbol{W}(\tau)\Big)$ is given by

$$\begin{pmatrix} \sum_{\ell=1}^{d}\sum_{k=1}^{d} \frac{\partial G_{11}}{\partial u_k} G_{k\ell}(\boldsymbol{u}) I_\ell(r,s) & \dots & \sum_{\ell=1}^{d}\sum_{k=1}^{d} \frac{\partial G_{1d}}{\partial u_k} G_{k\ell}(\boldsymbol{u}) I_\ell(r,s) \\ \vdots & \ddots & \vdots \\ \sum_{\ell=1}^{d}\sum_{k=1}^{d} \frac{\partial G_{d1}}{\partial u_k} G_{k\ell}(\boldsymbol{u}) I_\ell(r,s) & \dots & \sum_{\ell=1}^{d}\sum_{k=1}^{d} \frac{\partial G_{dd}}{\partial u_k} G_{k\ell}(\boldsymbol{u}) I_\ell(r,s) \end{pmatrix}.$$

By substituting the above, we finally get that the double integral is given by the *following expression*

$$\int_r^t DG(\boldsymbol{u}(r))\Big(G(\boldsymbol{u}(r))\int_r^s d\,\boldsymbol{W}(\tau)\Big)d\,\boldsymbol{W}(s) = \begin{pmatrix} \displaystyle\sum_{\ell=1}^d\sum_{k=1}^d\sum_{j=1}^d \frac{\partial G_{1j}}{\partial u_k} G_{k\ell}(\boldsymbol{u})I_j(r,t)I_\ell(r,s) \\ \vdots \\ \displaystyle\sum_{\ell=1}^d\sum_{k=1}^d\sum_{j=1}^d \frac{\partial G_{dj}}{\partial u_k} G_{k\ell}(\boldsymbol{u})I_j(r,t)I_\ell(r,s) \end{pmatrix}.$$

Then we expand the double integrals $I_j(r,t)I_\ell(r,s)$ in the following way. When $j=\ell$, we have that

$$I_j(r,t)I_\ell(r,s) = \frac{1}{2}\big(W_\ell(t) - W_\ell(r)\big)^2 - (t-r),$$

whereas for $j \neq \ell$,

$$I_j(r,t)I_\ell(r,s) = \int_r^t\int_r^s dW_\ell(\tau)dW_j(s) + \int_r^t\int_r^s dW_j(\tau)dW_\ell(s).$$

Finally, we define

$$A_{j\ell}(r,t) = \int_r^t\int_r^s dW_j(\tau)dW_\ell(s) - \int_r^t\int_r^s dW_\ell(\tau)dW_j(s).$$

Then the i-element of the double integral is given by

$$\frac{1}{2}\sum_{j=1}^d\sum_{k=1}^d \frac{\partial G_{ij}}{\partial u_k}(\boldsymbol{u})G_{kj}(\boldsymbol{u})\big(I_j^2(r,t) - \Delta t\big)$$

$$+\frac{1}{2}\sum_{j<\ell=1}^d\sum_{k=1}^d\left(\frac{\partial G_{i\ell}}{\partial u_k}(\boldsymbol{u})G_{kj}(\boldsymbol{u}) + \frac{\partial G_{ij}}{\partial u_k}(\boldsymbol{u})G_{k\ell}(\boldsymbol{u})\right)I_j(r,t)I_\ell(r,t)$$

$$+\frac{1}{2}\sum_{j<\ell=1}^d\sum_{k=1}^d\left(\frac{\partial G_{i\ell}}{\partial u_k}(\boldsymbol{u})G_{kj}(\boldsymbol{u}) - \frac{\partial G_{ij}}{\partial u_k}(bu)G_{k\ell}(\boldsymbol{u})\right)A_{j\ell}(r,t).$$

Example 4. *In the same context with Chapter 2 , we see in detail why the integral $\int_{t_n}^{t_{n+1}}\int_{t_n}^{t_{n+1}} DG(\boldsymbol{u}_n)G(\boldsymbol{u}_n)d\,\boldsymbol{W}(s)d\,\boldsymbol{W}(s)$ is equal to $\int_{t_n}^{t_{n+1}} \boldsymbol{c}(\boldsymbol{u}_n)ds$, with $\boldsymbol{c}(\boldsymbol{u}_n) \in \mathbb{R}^d$ as defined in Assumption 2.1.2.*

Firstly, we see that $G(\boldsymbol{u}_n)d\,\boldsymbol{W}$ is a vector in $\mathbb{R}^d$, since

$$\begin{pmatrix} G_{11}(\boldsymbol{u}_n) & \dots & G_{1d}(\boldsymbol{u}_n) \\ \vdots & \ddots & \vdots \\ G_{d1}(\boldsymbol{u}_n) & \dots & G_{dd}(\boldsymbol{u}_n) \end{pmatrix}\begin{pmatrix} dW_1 \\ \vdots \\ dW_d \end{pmatrix} = \begin{pmatrix} G_{11}(\boldsymbol{u}_n)dW_1 + \ldots + G_{1d}(\boldsymbol{u}_n)dW_d \\ \vdots \\ G_{d1}(\boldsymbol{u}_n)dW_1 + \ldots + G_{dd}(\boldsymbol{u}_n)dW_d \end{pmatrix}.$$

Then, $DG(\boldsymbol{u}_n)G(\boldsymbol{u}_n)d\,\boldsymbol{W}$ is a matrix in $\mathbb{R}^{d\times d}$, since we multiply the rank-three tensor

$DG(\boldsymbol{u}_n)$ by the vector $G(\boldsymbol{u}_n)d\boldsymbol{W}$ in the following way

$$\begin{pmatrix} \nabla G_{11}(\boldsymbol{u}_n) & \dots & \nabla G_{1d}(\boldsymbol{u}_n) \\ \vdots & \ddots & \vdots \\ \nabla G_{d1}(\boldsymbol{u}_n) & \dots & \nabla G_{dd}(\boldsymbol{u}_n) \end{pmatrix} \begin{pmatrix} G_{11}(\boldsymbol{u}_n)dW_1 + \ldots + G_{1d}(\boldsymbol{u}_n)dW_d \\ \vdots \\ G_{d1}(\boldsymbol{u}_n)dW_1 + \ldots + G_{dd}(\boldsymbol{u}_n)dW_d \end{pmatrix} =$$

$$\begin{pmatrix} \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{11}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j & \dots & \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{1d}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{d1}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j & \dots & \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{dd}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j \end{pmatrix}.$$

Finally, the product of the $d \times d$ matrix $DG(\boldsymbol{u}_n)G(\boldsymbol{u}_n)d\boldsymbol{W}$ and the vector $d\boldsymbol{W}$ gives us $\boldsymbol{c} \in \mathbb{R}^d$

$$\begin{pmatrix} \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{11}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j & \dots & \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{1d}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{d1}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j & \dots & \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{dd}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j \end{pmatrix} \begin{pmatrix} dW_1 \\ \vdots \\ dW_d \end{pmatrix} =$$

$$\begin{pmatrix} \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{1j}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j^2 \\ \vdots \\ \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{dj}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dW_j^2 \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{1j}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dt \\ \vdots \\ \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{dj}(\boldsymbol{u}_n)}{\partial u_k} G_{kj} dt \end{pmatrix} = \int_{t_n}^{t_{n+1}} \hat{\mathbf{c}}(\boldsymbol{u}_n) ds,$$

Note that, we have used the rule of Lemma 1.4.1 in order to substitute

$$\sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{ij}(\boldsymbol{u})}{\partial u_k} G_{kj}(\boldsymbol{u}) dW_j^2$$

by

$$\sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial G_{ij}(\boldsymbol{u})}{\partial u_k} G_{kj}(\boldsymbol{u}) dt.$$

Example 5. *This is an example where we compute the so called correction term of the Stratonovich interpreted Langevin system (2.20) in order to convert it into its equivalent Itô form. For the correction term we use the $\mathbf{c}(\mathbf{u})$ notation. So for the* *system (2.20), given the diffusion term*

$$G(\boldsymbol{u}) = \begin{pmatrix} 0 & -\sigma u_3 & \sigma u_2 \\ \sigma u_3 & 0 & -\sigma u_1 \\ -\sigma u_2 & \sigma u_1 & 0 \end{pmatrix},$$

we compute the terms $c_i(\boldsymbol{u})$ *for each* $i = 1, 2, 3$ *in the following way*

$$c_1(\boldsymbol{u}) = \sum_{k=1}^{3}\sum_{j=1}^{3} \frac{\partial G_{1j}(\boldsymbol{u})}{\partial u_k} G_{kj}(\boldsymbol{u}) = \frac{\partial G_{11}(\boldsymbol{u})}{\partial u_1} G_{11} + \frac{\partial G_{11}(\boldsymbol{u})}{\partial u_2} G_{21} + \frac{\partial G_{11}(\boldsymbol{u})}{\partial u_3} G_{31}$$

$$+\frac{\partial G_{12}(\boldsymbol{u})}{\partial u_1} G_{12} + \frac{\partial G_{12}(\boldsymbol{u})}{\partial u_2} G_{22} + \frac{\partial G_{12}(\boldsymbol{u})}{\partial u_3} G_{32} + \frac{\partial G_{13}(\boldsymbol{u})}{\partial u_1} G_{13} + \frac{\partial G_{13}(\boldsymbol{u})}{\partial u_2} G_{23} + \frac{\partial G_{13}(\boldsymbol{u})}{\partial u_3} G_{33}$$
$$= -2\sigma^2 u_1.$$

Similarly,

$$c_2(\boldsymbol{u}) = \sum_{k=1}^{3}\sum_{j=1}^{3} \frac{\partial G_{2j}(\boldsymbol{u})}{\partial u_k} G_{kj}(\boldsymbol{u}) = \frac{\partial G_{21}(\boldsymbol{u})}{\partial u_1} G_{11} + \frac{\partial G_{21}(\boldsymbol{u})}{\partial u_2} G_{21} + \frac{\partial G_{21}(\boldsymbol{u})}{\partial u_3} G_{31}$$

$$+\frac{\partial G_{22}(\boldsymbol{u})}{\partial u_1} G_{12} + \frac{\partial G_{22}(\boldsymbol{u})}{\partial u_2} G_{22} + \frac{\partial G_{22}(\boldsymbol{u})}{\partial u_3} G_{32} + \frac{\partial G_{23}(\boldsymbol{u})}{\partial u_1} G_{13} + \frac{\partial G_{23}(\boldsymbol{u})}{\partial u_2} G_{23} + \frac{\partial G_{23}(\boldsymbol{u})}{\partial u_3} G_{33}$$
$$= -2\sigma^2 u_2.$$

Finally,

$$c_3(\boldsymbol{u}) = \sum_{k=1}^{3}\sum_{j=1}^{3} \frac{\partial G_{3j}(\boldsymbol{u})}{\partial u_k} G_{kj}(\boldsymbol{u}) = \frac{\partial G_{31}(\boldsymbol{u})}{\partial u_1} G_{11} + \frac{\partial G_{31}(\boldsymbol{u})}{\partial u_2} G_{21} + \frac{\partial G_{31}(\boldsymbol{u})}{\partial u_3} G_{31}$$

$$+\frac{\partial G_{32}(\boldsymbol{u})}{\partial u_1} G_{12} + \frac{\partial G_{32}(\boldsymbol{u})}{\partial u_2} G_{22} + \frac{\partial G_{32}(\boldsymbol{u})}{\partial u_3} G_{32} + \frac{\partial G_{33}(\boldsymbol{u})}{\partial u_1} G_{13} + \frac{\partial G_{33}(\boldsymbol{u})}{\partial u_2} G_{23} + \frac{\partial G_{33}(\boldsymbol{u})}{\partial u_3} G_{33}$$
$$= -2\sigma^2 u_3.$$

Example 6.

In this example we expand the double integral from equation (2.22) of the MSEI2 Section 2.3.1.

$$\int_r^t e^{(t-s)A} \left(\frac{G(\boldsymbol{u}(r)) + G(\boldsymbol{u}(t))}{2} \int_{\frac{r+t}{2}}^{s} d\,\boldsymbol{W}(\tau) \right) \circ d\,\boldsymbol{W}(s) =$$

$$\int_r^t e^{(t-s)A} \frac{1}{2}\big(G(\boldsymbol{u}(r)) + G(\boldsymbol{u}(t))\big)\Big(\boldsymbol{W}(s) - \boldsymbol{W}(\frac{r+t}{2})\Big) \circ d\,\boldsymbol{W}(s) =$$

$$\frac{1}{2}\big(G(\boldsymbol{u}(r)) + G(\boldsymbol{u}(t))\big) \int_r^t e^{(t-s)A}\big(\boldsymbol{W}(s) - \boldsymbol{W}(\frac{r+t}{2})\big) \circ d\,\boldsymbol{W}(s) \approx$$

$$\frac{1}{2}\big(G(\boldsymbol{u}(r))+G(\boldsymbol{u}(t))\big)e^{\frac{(t-r)A}{2}}\Big(\int_r^t \boldsymbol{W}(s)\circ d\,\boldsymbol{W}(s)-\int_r^t \boldsymbol{W}(\frac{r+t}{2})\circ d\,\boldsymbol{W}(s)\Big).$$

We compute the two Stratonovich integrals

$$\int_r^t \boldsymbol{W}(s)\circ d\,\boldsymbol{W}(s)=\frac{1}{2}\big(\boldsymbol{W}^2(t)-\boldsymbol{W}^2(r)\big)$$

and

$$\int_r^t \boldsymbol{W}(\frac{r+t}{2})\circ d\,\boldsymbol{W}(s)=\boldsymbol{W}(\frac{r+t}{2})\big(\boldsymbol{W}(t)-\boldsymbol{W}(r)\big).$$

Then, we substitute for the integrals in the equation above, so

$$\int_r^t e^{(t-s)A}\left(\frac{G(\boldsymbol{u}(r))+G(\boldsymbol{u}(t))}{2}\int_{\frac{r+t}{2}}^s d\,\boldsymbol{W}(\tau)\right)\circ d\,\boldsymbol{W}(s)=$$

$$\frac{G(\boldsymbol{u}(r))+G(\boldsymbol{u}(t))}{2}e^{\frac{(t-r)A}{2}}\frac{1}{2}\Big(\boldsymbol{W}^2(t)-\boldsymbol{W}^2(r)-2\,\boldsymbol{W}(\frac{r+t}{2})\,\boldsymbol{W}(t)+2\,\boldsymbol{W}(\frac{r+t}{2})\,\boldsymbol{W}(r)\Big)=$$
$$\frac{G(\boldsymbol{u}(r))+G(\boldsymbol{u}(t))}{4}e^{\frac{(t-r)A}{2}}\Big(\big(\boldsymbol{W}(t)-\boldsymbol{W}(\frac{r+t}{2})\big)^2-\big(\boldsymbol{W}(r)-\boldsymbol{W}(\frac{r+t}{2})\big)^2\Big).$$

Example 7. *In this example we see how to obtain a component-wise expression of the double stochastic integral of equation (2.22).*

$$\int_r^t e^{(t-s)A}\frac{DG(\boldsymbol{u}(r))+DG(\boldsymbol{u}(t))}{2}\left(\frac{G(\boldsymbol{u}(r))+G(\boldsymbol{u}(t))}{2}\int_{\frac{r+t}{2}}^s d\,\boldsymbol{W}(\tau)\right)\circ d\,\boldsymbol{W}(s).$$

Unlike Example 6, in this example we also include the term $\frac{1}{2}\big(DG(\boldsymbol{u}(r))+DG(\boldsymbol{u}(t))\big)$*. This term does not depend on integration but is used here under the scope of performing tensors operations.*

Firstly, we simplify the notation by using $\tilde{G}(\boldsymbol{u})$ *instead of* $\frac{1}{2}\big(G(\boldsymbol{u}(r))+G(\boldsymbol{u}(t))\big)$ *and also by using* $\widetilde{DG}(\boldsymbol{u})$ *instead of* $\frac{1}{2}\big(DG(\boldsymbol{u}(r))+DG(\boldsymbol{u}(t))\big)$*. Moreover, we define* $I_i(\frac{r+t}{2},s):=\int_{\frac{r+t}{2}}^s dW_i(\tau)$ *and* $J_i(r,t):=\int_r^t \circ dW_i(s)$*. Then,*

$$\tilde{G}(\boldsymbol{u})\int_{\frac{r+t}{2}}^s d\,\boldsymbol{W}(\tau)=\begin{pmatrix} G_{11} & \dots & G_{1d}\\ \vdots & \ddots & \vdots\\ G_{d1} & \dots & G_{dd}\end{pmatrix}\begin{pmatrix}\int_{\frac{r+t}{2}}^s dW_1(\tau)\\ \vdots\\ \int_{\frac{r+t}{2}}^s dW_d(\tau)\end{pmatrix}$$

$$=\begin{pmatrix} G_{11}I_1(\frac{r+t}{2},s)+\ldots+G_{1d}I_d(\frac{r+t}{2},s)\\ \vdots\\ G_{d1}I_1(\frac{r+t}{2},s)+\ldots+G_{dd}I_d(\frac{r+t}{2},s)\end{pmatrix}:=\boldsymbol{b}$$

Next,

$$e^{\frac{\Delta t A}{2}}\widetilde{DG}(\boldsymbol{u})\boldsymbol{b} = \begin{pmatrix} \hat{a}_{11} & \dots & \hat{a}_{1d} \\ \vdots & \ddots & \vdots \\ \hat{a}_{d1} & \dots & \hat{a}_{dd} \end{pmatrix} \begin{pmatrix} \nabla G_{11}^\top \boldsymbol{b} & \dots & \nabla G_{1d}^\top \boldsymbol{b} \\ \vdots & \ddots & \vdots \\ \nabla G_{d1}^\top \boldsymbol{b} & \dots & \nabla G_{dd}^\top \boldsymbol{b} \end{pmatrix} =$$

$$\begin{pmatrix} \hat{a}_{11}\nabla G_{11}^\top \boldsymbol{b} + \ldots + \hat{a}_{1d}\nabla G_{d1}^\top \boldsymbol{b} & \dots & \hat{a}_{11}\nabla G_{1d}^\top + \ldots + \hat{a}_{1d}\nabla G_{dd}^\top \boldsymbol{b} \\ \vdots & \ddots & \vdots \\ \hat{a}_{d1}\nabla G_{11}^\top \boldsymbol{b} + \ldots + \hat{a}_{dd}\nabla G_{d1}^\top \boldsymbol{b} & \dots & \hat{a}_{d1}\nabla G_{1d}^\top \boldsymbol{b} + \ldots + \hat{a}_{dd}\nabla G_{dd}^\top \boldsymbol{b} \end{pmatrix},$$

where $\hat{a}_{ij}$, $i, j = 1, \ldots, d$ are the elements of the $e^{\frac{\Delta t A}{2}}$ operator. Finally,

$$\int_r^t e^{\frac{\Delta t A}{2}}\widetilde{DG}(\boldsymbol{u})\boldsymbol{b} \circ d\boldsymbol{W}(s) =$$

$$\begin{pmatrix} \big(\hat{a}_{11}\nabla G_{11}^\top \boldsymbol{b} + \ldots + \hat{a}_{1d}\nabla G_{d1}^\top \boldsymbol{b}\big)J_1(r,t) + \ldots + \big(\hat{a}_{11}\nabla G_{1d}^\top \boldsymbol{b} + \ldots + \hat{a}_{1d}\nabla G_{dd}^\top \boldsymbol{b}\big)J_d(r,t) \\ \vdots \\ \big(\hat{a}_{d1}\nabla G_{11}^\top \boldsymbol{b} + \ldots + \hat{a}_{dd}\nabla G_{d1}^\top \boldsymbol{b}\big)J_1(r,t) + \ldots + \big(\hat{a}_{d1}\nabla G_{1d}^\top \boldsymbol{b} + \ldots + \hat{a}_{dd}\nabla G_{dd}^\top\big)J_d(r,t) \end{pmatrix}.$$

If we further expand $\boldsymbol{b}$, the above is equivalent to

$$\begin{pmatrix} \sum_{q=1}^d \sum_{j=1}^d \hat{a}_{1q}\frac{\partial G_{qj}}{\partial u_1}G_{11}I_1J_j + \ldots + \hat{a}_{1q}\frac{\partial G_{qj}}{\partial u_1}G_{1d}I_dJ_j + \ldots + \hat{a}_{1q}\frac{\partial G_{qj}}{\partial u_d}G_{d1}I_1J_j + \ldots + \hat{a}_{1q}\frac{\partial G_{qj}}{\partial u_d}G_{dd}I_dJ_j \\ \vdots \\ \sum_{q=1}^d \sum_{j=1}^d \hat{a}_{dq}\frac{\partial G_{qj}}{\partial u_1}G_{11}I_1J_j + \ldots + \hat{a}_{dq}\frac{\partial G_{qj}}{\partial u_1}G_{1d}I_dJ_j + \ldots + \hat{a}_{dq}\frac{\partial G_{qj}}{\partial u_d}G_{d1}I_1J_j + \ldots + \hat{a}_{dq}\frac{\partial G_{qj}}{\partial u_d}G_{dd}I_dJ_j \end{pmatrix}.$$

So, after we replace $\tilde{G}(\boldsymbol{u})$ with $\frac{1}{2}\big(G(\boldsymbol{u}_n)+G(\tilde{\boldsymbol{u}}_{n+1})\big)$ and $\widetilde{DG}(\boldsymbol{u})$ with $\frac{1}{2}\big(DG(\boldsymbol{u}_n)+DG(\tilde{\boldsymbol{u}}_{n+1})\big)$, the k-th component of the double stochastic integral is given by the expression

$$\sum_{q=1}^d \sum_{j=1}^d \sum_{\ell=1}^d \hat{a}_{kq}\left(\frac{\partial G_{qj}}{\partial u_\ell}(\boldsymbol{u}_n)G_{\ell j}(\boldsymbol{u}_n) + \frac{\partial G_{qj}}{\partial u_\ell}(\tilde{\boldsymbol{u}}_{n+1})G_{\ell j}(\tilde{\boldsymbol{u}}_{n+1})\right)I_j(\frac{t_n+t_{n+1}}{2}, t_{n+1})J_j(t_n,t_{n+1})$$

$$+\frac{1}{8}\sum_{q=1}^d \sum_{j<i=1}^d \sum_{\ell=1}^d \hat{a}_{kq}\left[\left(\frac{\partial G_{qi}}{\partial u_\ell}(\boldsymbol{u}_n)G_{\ell j}(\boldsymbol{u}_n) + \frac{\partial G_{qi}}{\partial u_\ell}(\tilde{\boldsymbol{u}}_{n+1})G_{\ell j}(\tilde{\boldsymbol{u}}_{n+1})\right)\right.$$

$$\left.+\left(\frac{\partial G_{qj}}{\partial u_\ell}(\boldsymbol{u}_n)G_{\ell i}(\boldsymbol{u}_n) + \frac{\partial G_{qj}}{\partial u_\ell}(\tilde{\boldsymbol{u}}_{n+1})G_{\ell i}(\tilde{\boldsymbol{u}}_{n+1})\right)\right]I_i(\frac{t_n+t_{n+1}}{2}, t_{n+1})J_j(t_n,t_{n+1})$$

$$+\frac{1}{8}\sum_{q=1}^d \sum_{j<i=1}^d \sum_{\ell=1}^d \hat{a}_{kq}\left[\left(\frac{\partial G_{qi}}{\partial u_\ell}(\boldsymbol{u}_n)G_{\ell j}(\boldsymbol{u}_n) + \frac{\partial G_{qi}}{\partial u_\ell}(\tilde{\boldsymbol{u}}_{n+1})G_{\ell j}(\tilde{\boldsymbol{u}}_{n+1})\right)\right.$$

$$\left.\times\left(\frac{\partial G_{qj}}{\partial u_\ell}(\boldsymbol{u}_n)G_{\ell i}(\boldsymbol{u}_n) + \frac{\partial G_{qj}}{\partial u_\ell}(\tilde{\boldsymbol{u}}_{n+1})G_{\ell i}(\tilde{\boldsymbol{u}}_{n+1})\right)\right]A_{ij,n}. \tag{7.1}$$

7.2 Some proofs from Chapters 1 and 3

The following is a sketch of the proof for Lemma 1.4.1 where we follow the approach of [61].

Proof.
Equivalently with $\mathbb{E}\big((dW)(t)\big)^2 = dt$ we can prove that

$$\lim_{\Delta t \to 0} \sum_{k=0}^{n-1} u(t_k)\Delta W_k^2 = \lim_{\Delta t \to 0} \sum_{k=0}^{n-1} u(t_k)\Delta t,$$

where the limit is taken in the mean square sense. Thus

$$\int_0^t u(s)dW^2(s) = \int_0^t u(s)ds.$$

In order to show that, we consider the following difference

$$\mathbb{E}\Big(\big(\sum_{k=0}^{n-1} u(t_k)\Delta W_k^2 - \sum_{k=0}^{n-1} u(t_k)\Delta t\big)^2\Big) = \mathbb{E}\Big(\big(\sum_{k=0}^{n-1} u(t_k)(\Delta W_k^2 - \Delta t)\big)^2\Big)$$

$$= \sum_{k=0}^{n-1}\sum_{\ell=0}^{n-1} \mathbb{E}\Big(u(t_k)u(t_\ell)\big(\Delta W_k^2 - \Delta t\big)\big(\Delta W_\ell^2 - \Delta t\big)\Big).$$

Note that when $k = \ell$, indeed

$$u(t_k)u(t_\ell) = u(t_k)^2 \text{ and } \big((\Delta W_k^2 - \Delta t)\big)^2 = \big(\Delta W_k^2 - \Delta t\big)\big(\Delta W_k^2 - \Delta t\big)\Big).$$

Otherwise, if $k > \ell$, then

$$\big((\Delta W_k^2 - \Delta t)\big) \text{ is independent of } u(t_k)u(t_\ell)\big(\Delta W_\ell^2 - \Delta t\big),$$

so

$$\mathbb{E}\Big(u(t_k)u(t_\ell)\big(\Delta W_k^2 - \Delta t\big)\big(\Delta W_\ell^2 - \Delta t\big)\Big) =$$

$$\mathbb{E}\Big(u(t_k)u(t_\ell)\big(\Delta W_\ell^2 - \Delta t\big)\Big)\mathbb{E}\Big(\Delta W_k^2 - \Delta t\Big) = 0.$$

Similarly if $k < \ell$, then

$$\big((\Delta W_\ell^2 - \Delta t)\big) \text{ is independent of } u(t_k)u(t_\ell)\big(\Delta W_k^2 - \Delta t\big),$$

so

$$\mathbb{E}\Big(u(t_k)u(t_\ell)\big(\Delta W_k^2 - \Delta t\big)\big(\Delta W_\ell^2 - \Delta t\big)\Big) =$$

$$\mathbb{E}\Big(u(t_k)u(t_\ell)\big(\Delta W_k^2 - \Delta t\big)\Big)\mathbb{E}\Big(\Delta W_\ell^2 - \Delta t\Big) = 0.$$

Thus,

$$\sum_{k=0}^{n-1}\sum_{\ell=0}^{n-1}\mathbb{E}\Big(u(t_k)u(t_\ell)\big(\Delta W_k^2-\Delta t\big)\big(\Delta W_\ell^2-\Delta t\big)\Big) = \sum_{k=0}^{n-1}\mathbb{E}\Big(u(t_k)^2\big(\Delta W_k^2-\Delta t\big)^2\Big) \tag{7.2}$$

$$= \sum_{k=0}^{n-1}\mathbb{E}\Big(u(t_k)^2\Big)\mathbb{E}\Big(\big(\Delta W_k^2 - \Delta t\big)^2\Big),$$

since ΔW_k is independent of $u(t_k)$. Next, we are going to use the following property

$$\mathbb{E}\Big(\big(W(t) - W(s)\big)^2\Big) = \mathrm{Var}\Big(W(t) - W(s)\Big) = t - s. \tag{7.3}$$

So,

$$\mathbb{E}\Big(\big(\Delta W_k^2 - \Delta t\big)^2\Big) = \mathrm{Var}\big(\Delta W_k^2 - \Delta t\big).$$

Also, since $\mathrm{Var}\big(X + \alpha\big) = \mathrm{Var}\big(X\big)$, $\alpha \in \mathbb{R}$,

$$\mathbb{E}\Big(\big(\Delta W_k^2 - \Delta t\big)^2\Big) = \mathrm{Var}\big(\Delta W_k^2\big).$$

By (7.3),

$$\mathrm{Var}\big(\Delta W_k^2\big) = \mathrm{Var}\Big(\big(W(t_k + \Delta t) - W(t_k)\big)^2\Big) = \Delta t^2,$$

hence,

$$\mathbb{E}\Big(\big(\Delta W_k^2 - \Delta t\big)^2\Big) = \Delta t^2.$$

Then (7.2) becomes

$$\mathbb{E}\Bigg(\Big(\sum_{k=0}^{n-1} u(t_k)\Delta W_k^2 - \sum_{k=0}^{n-1} u(t_k)\Delta t\Big)^2\Bigg) = \sum_{k=0}^{n-1}\mathbb{E}\Big(u(t_k)^2\Big)\Delta t^2.$$

Note that $\sum_{k=0}^{n-1}\mathbb{E}\Big(u(t_k)^2\Big)\Delta t^2 \to 0$, as $\Delta t \to 0$, thus

$$\lim_{\Delta t\to 0}\sum_{k=0}^{n-1} u(t_k)\Delta W_k^2 = \lim_{\Delta t\to 0}\sum_{k=0}^{n-1} u(t_k)\Delta t,$$

which implies

$$\int_0^t u(s)dW^2(s) = \int_0^t u(s)ds.$$

$\square$

The following is a sketch of the proof for Proposition 3.1.2 for $d = m = 1$.

Proof. We start from looking at the difference

$$\phi\big(t_j,(1-\alpha)u_j+\alpha u_{j+1}\big)-\phi(t_j,u_j).$$

To simplify the notation, we use $\phi(u_j)$ instead of $\phi(t_j,u(t_j))$, where u_j denotes $u(t_j)$, $t_j=j\Delta t$, $j=0,\dots,N-1$. From the mean value theorem we have that there exists some $\theta_j\in(0,1)$ such that

$$\phi\big((1-\alpha)u_j+\alpha u_{j+1}\big)-\phi(u_j)=\alpha\phi'\big((1-\alpha)(\frac{1}{\alpha}-\theta_j)u_j+\alpha\theta_j u_{j+1}\big)(u_{j+1}-u_j).$$

Since u satisfies equation (3.1), the above equality is equivalent to

$$\phi\big((1-\alpha)u_j+\alpha u_{j+1}\big)-\phi(u_j)=\alpha\phi'\big((1-\alpha)(\frac{1}{\alpha}-\theta_j)u_j+\alpha\theta_j u_{j+1}\big)$$

$$\times\big(F(u_j)\Delta t_j+G(u_j)\Delta W_j+\text{h.o.t}\big), \tag{7.4}$$

where $\Delta t_j=t_{j+1}-t_j$ and $\Delta W_j=W(t_{j+1})-W(t_j)$. We multiply both sides of (7.4) by ΔW_j and we set $\rho:=(1-\alpha)(\frac{1}{\alpha}-\theta_j)u_j+\alpha\theta_j u_{j+1}$. Then,

$$\phi(\rho)\Delta W_j=\phi(u_j)\Delta W_j+\alpha\phi'(\rho)F(u_j)\Delta t_j\Delta W_j+\alpha\phi'(\rho)G(u_j)\Delta W_j^2+\text{h.o.t.}$$

By considering the expected value of both sides and based on the fact that $\mathbb{E}[\Delta t_j\Delta W_j]=0$ and $\mathbb{E}[\Delta W_j^2]=\Delta t_j$, we get

$$\mathbb{E}[\phi(\rho)\Delta W_j]=\mathbb{E}[\phi(u_j)\Delta W_j+\alpha\phi'(\rho)G(u_j)\Delta t_j]. \tag{7.5}$$

So, by summing up the terms of (7.5) and taking limits in both sides we finally have

$$\mathbb{E}\Big[\lim_{N\to\infty}\sum_{j=0}^{N-1}\phi\big((1-\alpha)(\frac{1}{\alpha}-\theta_j)u_j+\alpha\theta_j u_{j+1}\big)\Delta W_j\Big]=\mathbb{E}\Big[\lim_{N\to\infty}\sum_{j=0}^{N-1}\phi(u_j)\Delta W_j\Big]$$

$$+\mathbb{E}\big[\lim_{N\to\infty}\sum_{j=0}^{N-1}\alpha\phi'\big((1-\alpha)(\frac{1}{\alpha}-\theta_j)u_j+\alpha\theta_j u_{j+1}\big)G(u_j)\Delta t_j\big],$$

or

$$\int_0^t\phi(u(s))*dW(s)=\int_0^t\phi(u(s))dW(s)+\alpha\int_0^t\phi'(u(s))G(u(s))ds. \tag{7.6}$$

□

Note that this generalisation uses as a reference the Itô integral, this basically means that it gives a formula which connects a general stochastic integral with the Itô integral. In a similar way, we can use the Stratonovich integral as a reference to obtain the following correction formula

$$\int_0^t \phi(u(s)) * dW(s) = \int_0^t \phi(u(s)) \circ dW(s) + (\alpha - \frac{1}{2}) \int_0^t \phi'(u(s)) G(u(s)) ds.$$

7.3 Matlab codes

Algorithm 7.3.1 Euler-Maruyama function file.

```
%%% SODE du=f(u)dt+g(u)dW(t).
%%% function with inputs:
%%%         u0: initial condition
%%%         T : time interval [0,T]
%%%         N : number of iterations
%%%         d : order of the ODE (f:R^d->R^d)
%%%         m : dimension of the Wiener process
%%% fhandle : drift term
%%% ghandle : diffusion term.
%%% The outputs are a time vector t, approximate solution u.
function [t, u]=EulerMaruyama2(u0, T, N, d, m, fhandle, ghandle )

Dt=T/N;                 %%% step-size
u=zeros(d,N+1);
t=zeros(N+1,1);
u_n=u0;                 %%% preallocate u
sqrtDt=sqrt(Dt);
for n=1:N+1
u(:,n)=u_n; t(n)=(n-1)*Dt;
dW=sqrtDt*randn(m,1);                          %%% Brownian increments
u_new=u_n+Dt*fhandle(u_n)+ghandle(u_n)*dW;     %%% E-M scheme
u_n=u_new;
end
end
```

Algorithm 7.3.2 Strong convergence for Euler-Murayama.

```
clear all
fhandle=@(x) x; ghandle=@(x) 0.1*x; dghandle=@(x) 0.1; hhandle=@(x) x;
M=200;                                                  % number of samples
u0=ones(1,M);                                           % initial condition
DT=[]; error={}; normerror=[]; errortemp=[];
T=1; N=2^14; d=1; m=1;
%% Generates reference solution
tic;
randn('state',100);
[u,t]=EMpath( u0,T,N,d,m,fhandle,ghandle,1,M);
uref=squeeze(u);
toc
%% Generates paths with bigger time step-sizes
udt={};
tic
kappa=[2 2^2 2^3 2^4 2^5 2^6];
for i=1:length(kappa);
randn('state',100);
xr(i)=cputime;

[u,t]=EMpath( u0,T,N,d,m,fhandle,ghandle,kappa(i),M);

udt{i}=squeeze(u);
Dt(i)=t(end)-t(end-1);
error{i}=abs(uref(:,end)-udt{i}(:,end));
errortemp=error{i};
normerror(i)=sqrt(mean(errortemp)/M);

cputime-xr(i);
end
toc

xles=log(Dt); yles=log(normerror); fl=polyfit(xles,yles,1);
fprintf('slope of the fitted line is %f \n',fl(1))

loglog(Dt,normerror,'-*')
hold on
loglog(Dt,0.1*Dt.^(0.5),'r-')
```

Algorithm 7.3.3 Function that generates M samples of E-M paths. Each path is generated by using several sizes of Δt, the smallest of which is considered as the reference Δt.

```
function [u,t] = EMpath( u0,T,N,d,m,fhandle,ghandle,kappa,M)
Dtref=T/N;
Dt=kappa*Dtref;
NN=N/kappa;
u=zeros(d,M,NN+1);
gdW=zeros(d,M);
sqrtDtref=sqrt(Dtref);
u_n=u0;
for n=1:NN+1
    t(n)=(n-1)*Dt;
    u(:,:,n)=u_n;
    dW=sqrtDtref*squeeze(sum(randn(m,M,kappa),3));
    for mm=1:M
        gdW(:,mm)=ghandle(u_n(:,mm))*dW(:,mm);
    end
    u_new=u_n+Dt*fhandle(u_n)+gdW;
    u_n=u_new;
end
```

Algorithm 7.3.4 Heun function file.

```

%%% STRATONOVICH
%%% SODE du=f(u)dt+g(u)o dW(t).
%%% function with inputs:
%%%         u0: initial condition
%%%         T : time interval [0,T]
%%%         N : number of iterations
%%%         d : order of the ODE (f:R^d->R^d)
%%%         m : dimension of the Wiener process
%%% fhandle : drift term
%%% ghandle : diffusion term.
%%% The outputs are a time vector t, approximate solution u and Dt.

function [t, u, Dt]=heun(u0, T, N, d, m, fhandle, ghandle)

Dt=T/N; u=zeros(d,N+1); t=zeros(N+1,1);
u_n=u0;
sqrtDt=sqrt(Dt);

for n=1:N+1
    u(:,n)=u_n;
    t(n)=(n-1)*Dt; dW=sqrtDt*randn(m,1);

    U_j=u_n+Dt*fhandle(u_n)+ghandle(u_n)*dW;
% E-M prediction of u(tilde)
    u_new=u_n+Dt*fhandle(u_n)+(0.5*(ghandle(u_n)+ghandle(U_j))*dW); %Heun
    u_n=u_new;
end
```

Algorithm 7.3.5 Exponential Integrators function file.

```
%% ITO
%% SODE du=f(u)dt+g(u)dW(t).
%% function with inputs:
%%          u0: initial condition
%%          T : time interval [0,T]
%%          N : number of iterations
%%          d : order of the ODE (f:R^d->R^d)
%%          m : dimension of the Wiener process
%% fhandle : drift term
%% ghandle : diffusion term.
%%          A : matrix(parameter)
%% The outputs are a time vector t, approximate solution u and Dt.

function [t, u, Dt]=Expon_Ito(u0, T, N, d, m, fhandle, ghandle, A)
Dt=T/N; u=zeros(N+1,d); t=zeros(N+1,1); u_n=u0;
sqrtDt=sqrt(Dt);

for n=1:N+1
   u(n,:)=u_n; t(n)=(n-1)*Dt; dW=sqrtDt*randn(m,1);
   u_new=exp(Dt*A)*u_n+Dt*fhandle(u_n)*((exp(Dt*A)-1)/(Dt*A))+...%exp integr
       exp(Dt*A)*ghandle(u_n)*dW;
   u_n=u_new;


end
```

Algorithm 7.3.6 Stratonovich Exponential Integrators (SEI) function file.

```
%%% STRATONOVICH
%%% SODE du=(Au+f(u))dt+g(u)o dW(t).
%%% function with inputs:
%%%        u0: initial condition
%%%        T : time interval [0,T]
%%%        N : number of iterations
%%%        d : order of the ODE (f:R^d->R^d)
%%%        m : dimension of the Wiener process
%%% fhandle : drift term
%%% ghandle : diffusion term.
%%%        A : bounded operator (parameter in 1-D)
%%% The outputs are a time vector t, approximate solution u and Dt.


function [t, ue, Dt]=Expon_Stratonovich(u0, T, N, d, m, fhandle, ghandle, A)

Dt =T/N ;
ue=zeros (d,N+1);
t=zeros (N+1,1);
ue_n=u0;
sqrtDt=sqrt(Dt);
phi0=expm(Dt*A);
phi05=expm(0.5*Dt*A);
phi1=A\(phi0-eye(size(A)));
for n=1:N+1
ue(:,n)=ue_n;
t(n)=(n-1)*Dt; dW=sqrtDt*randn(m,1);

%v_k=ue_n+(Dt*(A*(ue_n)+fhandle(ue_n)))+(ghandle(ue_n)*dW);  % E-M predictor
v_k= phi0*ue_n+(phi1*fhandle(ue_n))+(phi0*ghandle(ue_n)*dW); % exp predictor

ue_j=phi0*ue_n+phi1*fhandle(ue_n)+0.5*phi05*(ghandle(ue_n)+ghandle(v_k))*dW;

ue_n=ue_j;

end

plot(t,ue,'b')
xlabel('t ','FontSize',14)
ylabel('u(t)','FontSize',14 ,'Rotation',0)
```

Algorithm 7.3.7 SEI scheme as sample paths generator in 3-D.

```
function [t,u]=expo_sample_paths3(u0,T,N,d,m,fhandle,ghandle,kappa,M,A)

Dtref=T/N;
Dt=kappa*Dtref;
NN=N/kappa;
u=zeros(d,M,NN+1);
sqrtDtref=sqrt(Dtref);

u_n=u0;
gdW=zeros(d,M);
v_k=zeros(d,M);
phi0=expm(Dt*A);
phi05=expm(0.5*Dt*A);
phi1=A\(phi0-eye(size(A)));

for n=1:NN+1
    t(n)=(n-1)*Dt;      % for the big steps
    u(:,:,n)=u_n;

    dW=sqrtDtref*squeeze(sum(randn(m,M,kappa),3));

    for mm=1:M          % begin samples
        v_k(:,mm)=u_n(:,mm)+Dt*(A*(u_n(:,mm))+fhandle(u_n(:,mm)))+...
            ghandle(u_n(:,mm))*dW(:,mm);          %predictor for exp
        gdW(:,mm)=(0.5*phi05*(ghandle(u_n(:,mm))+ghandle(v_k(:,mm)))*dW(:,mm));



    end
    u_new=phi0*u_n+(phi1*fhandle(u_n))+gdW;

    u_n=u_new;
end
```

Algorithm 7.3.8 Strong convergence for SEI in higher dimensions.

```
%% STRONG CONVERGENCE FOR SEI IN 3-DIMENSIONS
clear all;
T=1; d=3; m=3; N=2^10; nu=0.1;
fhandle=@(x) 0.2*x;
ghandle=@(x) [0.2*x(1) 0 0; 0 -0.2*x(2) 0; 0 0 0.2*x(3)];
A=[1 0 0 ; 0 -1 0; 0 0 1];

M=100; u0=repmat([1;1;1],[1,M]);
DT=[];
error={};
errortemp=[];
%% GENERATES REFERENCE SOLUTION GIVEN BY SEI (U_EXACT~U_REF), M-SAMPLES
tic;
randn('state',100);
[t,u]=expo_sample_paths3(u0,T,N,d,m,fhandle,ghandle,1,M,A);
uref=u;
toc

%% GENERATES SOLUTIONS WITH DIFFERENT STEP-SIZES (DT_i=KAPPA(i)*DT),M-SAM
utemp=[]; udt={}; k=1;
tic
errormat=[];
kappa=[4 2^3 2^4 2^5 2^6 2^7];
for i=1:length(kappa);
    randn('state',100);
    [t,u]=expo_sample_paths3(u0,T,N,d,m,fhandle,ghandle,kappa(i),M,A);
    udt{i}=u; %squeeze(u);
    Dt(i)=t(end)-t(end-1);
    error{i}=uref(:,:,end)-udt{i}(:,:,end);
    utemp=uref(:,:,1:kappa(i):end);

    for j=1:M
    errormat(:,i)= error{i}(1,:).^2+error{i}(2,:).^2; % vector norm

    end

end
toc
normerror=sqrt(sum(errormat)/M);
%% PLOTS
loglog(Dt,normerror,'bo-')
hold on
loglog(Dt,0.6*abs(log(Dt)).*Dt,'-.')
hold on
loglog(Dt,0.1*Dt,'-.')
grid on
```

Algorithm 7.3.9 Milstein-type Stratonovich Exponential Integrators (MSEI) function file.

```
%%% STRATONOVICH EXPONENTIAL MILSTEIN
%%% SODE du=(Au+f(u))dt+g(u)o dW(t).
%%% function with inputs:
%%%        u0: initial condition
%%%        T : time interval [0,T]
%%%        N : number of iterations
%%%        d : order of the ODE (f:R^d->R^d)
%%%        m : dimension of the Wiener process
%%% fhandle : drift term
%%% ghandle : diffusion term.
%%%        A : matrix (dxd)
%%% The outputs are a time vector t, approximate solution u and Dt.

function [t,ue,Dt]=Exponential_Mil_dm(u0,T,N,d,m,fhandle,ghandle,dghandle,A)

Dt =T/N ;
ue=zeros (d,N+1); t=zeros (N+1,1); ue_n=u0;  sqrtDt=sqrt(Dt);

phi0=expm(Dt*A);
phi05=expm(0.5*Dt*A);
phi1=A\(phi0-eye(size(A)));
y2=0;

for n=1:N+1
ue(:,n)=ue_n;
t(n)=(n-1)*Dt; dW=sqrtDt*randn(m,1);

y1=0; Winc1=0; Winc2=0; Winc3=0;
dW1=dW(1); dW2=dW(2); dW3=dW(3);

y1=y1+y2*dW1;                % auxiliary system
y2=y2+dW2;                   % auxiliary system
Winc1=Winc1+dW1;
Winc2=Winc2+dW2;
Winc3=Winc3+dW3;

%v_k=ue_n+(Dt*(A*(ue_n)+fhandle(ue_n)))+(ghandle(ue_n)*dW); % E-M predictor
v_k= phi0*ue_n+(phi1*fhandle(ue_n(1),ue_n(2),ue_n(3)))+...
     (phi0*ghandle(ue_n(1),ue_n(2),ue_n(3))*dW);              % exp predictor

ue_j=phi0*ue_n+phi1*fhandle(ue_n(1),ue_n(2),ue_n(3))+...
    0.5*phi05*(ghandle(ue_n(1),ue_n(2),ue_n(3))+...
    ghandle(v_k(1),v_k(2),v_k(3)))*dW+...
    0.5*phi05*dghandle(ue_n(1),ue_n(2),ue_n(3)).*(dW.^2-Dt)+...
    0.5*phi05*dghandle(ue_n(1),ue_n(2),ue_n(3)).*...
    ([Winc1*Winc2;Winc1*Winc3;Winc2*Winc3]+[y1; y1; y1]);

ue_n=ue_j;
end
```

Bibliography

[1] I. A. Adamu, *Numerical Approximation of SDEs and the Stochastic Swift-Hohenberg Equation.* PhD Thesis, Heriot-Watt University, 2010.

[2] E. Alòs, D. Nualart, F. Viens, *Stochastic heat equation with white-noise drift.* Ann. Inst. H. Poincaré Probab. Statist., Vol. 36, no. 2, pp. 181-218, 2000.

[3] D. Applebaum, *Lévy processes and stochastic calculus.* 2nd Edition, Cambridge Studies in Advanced Mathematics 116, Cambridge University Press, Cambridge 2009.

[4] L. Arnold, *Stochastic differential equations: theory and applications.* Wiley-Interscience, John Wiley & Sons, 1974.

[5] L. Banas, Z. Brzezniak, A. Prohl, *Computational studies for the stochastic Landau-Lifshitz-Gilbert equation.* SIAM J. Sci. Comput., Vol. 35, no. 1, pp. B62-B81, 2013.

[6] N. Berglund, B. Gentz. Stochastic dynamic bifurcations and excitability. In C. Laing, G. J. Lord, editors, *Stochastic methods in Neuroscience.* Oxford University Press, Oxford, 2010.

[7] W.-J. Beyn, V. Thümmler, *Freezing solutions of equivariant evolution equations.* SIAM J. Appl. Dyn. Syst., Vol. 3, no. 2, pp. 85-116, 2004.

[8] S. C. Brenner, L. R. Scott, *The mathematical theory of finite element methods.* Springer, 2008.

[9] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations.* Springer, New York 2011.

[10] Z. Brzezniak, B. Goldys, T. Jegaraj, *Weak solutions of the stochastic Landau Lifschitz Gilbert equation.* Appl. Math. Res. Express. AMRX, no. 1, pp. 1-33, 2013.

[11] P. Chow, *Stochastic partial differential equations.* Chapman & Hall/CRC Applied Mathematics and Nonlinear Sciences Series. Chapman & Hall/CRC, Boca Raton, FL, 2007.

[12] S. M. Cox, P. C. Matthews, *Exponential time differencing for stiff systems.* J. Comp. Phys., Vol. 176, no. 2, pp. 430-455, 2002.

[13] G. Da Prato, J. Zabczyk, *Stochastic equations in infinite dimensions.* Cambridge University Press, Cambridge, 1992.

[14] L. C. Evans, *An introduction to stochastic differential equations.* Lecture notes Version 1.2, UC Berkeley 2006.

[15] G. W. Evans, S. Honkapohja, *Learning and expectations in macroeconomics.* Princeton University Press, New Jersey, 2001.

[16] J. García-Ojalvo, J. M.Sancho, *Noise in spatially extended systems.* Springer-Verlag, New-York, 1999.

[17] D. Givon, R. Kupferman, *White noise limits for discrete dynamical systems driven by fast deterministic dynamics.* Phys. A 335, no. 3-4, pp. 385-412, 2004.

[18] G. Gottwald, I. Melbourne, *Homogenization for deterministic maps and multiplicative noise.* Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 469, no. 2156, 2013.

[19] I. Gyöngy, *Existence and uniqueness results for semilinear stochastic partial differential equations.* Stochastic Process. Appl., Vol. 73, no. 2, pp. 271-299, 1998.

[20] D. Henry, *Geometric theory of semilinear parabolic equations: Lecture notes in mathematics, 840.* Springer-Verlag, 1981.

[21] D. J. Higham, *An algorithmic introduction to numerical simulation of stochastic differential equations.* SIAM Rev., Vol. 43, no. 3, pp. 525-546, 2001.

[22] D. J.Higham, P. E.Kloeden, *An introduction to the numerical simulation of stochastic differential equations.* Lecture notes for the Scottish Mathematical Sciences Training Centre, January 2008.

[23] D. J. Higham, P.E. Kloeden, *MAPLE and MATLAB for stochastic differential equations in finance, in: Programming languages and systems in computational economics and finance.* (S.S. Nielsen, ed.), Kluwer Academic Publishers, Springer, pp.233-270, 2002.

[24] D. J. Higham, X. Mao, A. M. Stuart, *Strong convergence of Euler-type methods for nonlinear stochastic differential equations.* SIAM J. Numer. Anal., Vol. 40, no. 3, pp.1041-1063, 2002.

[25] D. J. Higham, X. Mao, L. Szpruch, *Convergence, non-negativity and stability of a new Milstein-scheme with applications to finance.* Discrete Contin. Dyn. Syst. Ser B, Vol. 18, no. 8, pp.2083-2100, 2013.

[26] M. Hochbruck, C. Lubich, H. Selhofer, *Exponential integrators for large systems of differential equations.* SIAM J. Sci. Comput., Vol. 19, no. 5, pp. 1552-1574, 1998.

[27] M. Hochbruck, A. Ostermann, *Explicit exponential Runge-Kutta methods for semilinear parabolic problems.* SIAM J. Numer. Anal., Vol. 43, no. 3, pp. 1069-1090 , 2005.

[28] M. Hochbruck, A. Ostermann, *Exponential Runge-Kutta methods for parabolic problems.* Applied Numerical Mathematics, Vol. 53, no. 2-4, pp. 323-339 , 2005.

[29] M. Hochbruck, A. Ostermann, *Exponential Integrators.* Acta Numerica, Vol. 19, pp. 209-286, 2010.

[30] M. Hutzenthaler, A. Jentzen *Numerical approximations of stochastic differential equations with nonglobally Lipschitz continuous coefficients.* arXiv: 1203.5809v2.

[31] M. Hutzenthaler, A. Jentzen, P. E. Kloeden *Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients.* Ann. Appl. Probab., Vol. 22, no. 4, pp. 1611-1641, 2012.

[32] M. Hutzenthaler, A. Jentzen, E. Platen *Strong and weak divergence in finite time of Euler's method for SDEs with non-globally Lipschitz continuous coefficients.* Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 467, no.2130, pp. 1563-1576, 2011.

[33] S. Ilie, A. Teslya, *An adaptive step size method for the chemical Langevin equation.* J. Chem. Phys., 136, 2012.

[34] K. Itô, *On stochastic differential equations.* Mem. Amer. Math. Soc., no.4, 1951.

[35] E. M. Izhikevich, *Dynamical systems in Neuroscience: the geometry of excitability and bursting.* MIT Press, Cambridge, MA, 2007.

[36] A. Jentzen, P. E. Kloeden, *Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time*

noise. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., no. 2102, 465: 649-667, 2009.

[37] A. Jentzen, M. Röckner, *A Milstein scheme for SPDEs*. Found. Comput. Math., Vol. 15, no.2, pp. 313-362, 2015.

[38] C. Johnson, *Numerical solution of partial differential equations by the finite element method*. Cambridge University Press, Cambridge, 1987.

[39] I. Karatzas, S. Shreve, *Brownian motion and stochastic calculus.* Springer, New York, 1991.

[40] P. E. Kloeden, *The systematic derivation of higher order numerical schemes for stochastic differential equations*. Milan Journal of Mathematics, Vol. 70, pp. 187-207, 2002.

[41] P. E. Kloeden, G. J. Lord, A. Neuenkirch, T. Shardlow *The exponential integrator scheme for stochastic partial differential equations: pathwise error bounds.* J. Comput. Appl. Math., 235, no. 5, pp. 1245-1260, 2011.

[42] P. E. Kloeden, E. Platen, *Numerical solution of stochastic differential equations.* Springer-Verlag, Berlin, 1992.

[43] P. E. Kloeden, E. Platen, H. Schurz, *Numerical solution of SDE through computer experiments.* Springer-Verlag, Berlin, 1994.

[44] P. E. Kloeden, E. Platen, W. Wright, *The approximation of multiple stochastic integrals*. Stochastic Anal. Appl., Vol. 10, no.4, pp.431-441, 1992.

[45] A. Kohatsu-Higa, J. A León, *Anticipating stochastic differential equations of Stratonovich type*. Applied Mathematics and Optimization, Vol. 36, no.3, pp. 263-289, 1997.

[46] P. Kotelenez, *Stochastic ordinary and stochastic partial differential equations. Transition from microscopic to macroscopic equations.* Sringer, New York, 2008.

[47] N. V. Krylov, B. L. Rozovskii, *Itô equations in Banach spaces and strongly parabolic stochastic partial differential equations*. Dokl. Akad. Nauk. SSSR, 249, no. 2, pp. 285-289, 1979.

[48] R. Kupferman, G. A. Pavliotis, A. M. Stuart, *Itô versus Stratonovich white noise limits for systems with inertia and colored multiplicative noise*. Phys. Rev. E. (3), 70, no.3, 2004.

[49] D. Küpper, A. Kværnø, A. Rößler, *A Runge-Kutta method for index 1stochastic differential- algebraic equations with scalar noise.* BIT 52, no. 2, pp. 437-455, 2012.

[50] T. G. Kurtz, E. Pardoux, P. Protter, *Stratonovich stochastic differential equations driven by general semimartingales.* Ann. Inst. H. Poincaré, Statist., Vol. 31, no.2, pp. 351-377, 1995.

[51] S. Larsson, V. Thomée, *Partial differential equations with numerical methods.* Sringer, 2003.

[52] M. Ledoux, T. Lyons, Z. Qian, *Lévy area of Wiener processes in Banach spaces.* Ann. Probab., Vol. 30, no. 2, pp. 546-578, 2002.

[53] A. Lejay, *An introduction to rough paths.* Séminaire de Probabilités XXXVII, 1-59, Lecture Notes in Math., 1832, Springer, Berlin, 2003.

[54] G. J. Lord, C. E. Powell, T. Shardlow, *An introduction to computational stochastic PDEs.* Cambridge texts in applied mathematics, 2014.

[55] G. J. Lord, J. Rougemont, *A numerical scheme for stochastic PDEs with Gevrey regularity.* IMA J. Numer. Anal., Vol. 24, no. 4, pp. 587-604, 2004.

[56] G. J. Lord, A. Tambue, *Stochastic exponential integrators for the finite element discretization of SPDEsfor multiplicative and additive noise.* IMA J. Numer. Anal., Vol. 33, no. 2, pp. 515-543, 2013.

[57] G. J. Lord, V. Thümmler, *Computing stochastic traveling waves.* SIAM J. Sci. Comput., Vol. 34, no. 1, pp. B24-B43, 2012.

[58] X. Mao, *Stochastic differential equations and applications.* 2nd Edition, Horwood Publishing Limited, Chichester, 2008.

[59] X. Mao, L. Szpruch *Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients.* J. Comput. Appl. Math., 238, pp. 14-28, 2013.

[60] V. Mehrmann, P. Kunkel, *Differential-algebraic equations. Analysis and numerical solution.* EMS Textbooks in Mathematics, European Mathematical Society, Zürich, 2006.

[61] J. R. Movellan *Tutorial on Stochastic Differential Equations,* MPLab Tutorials, Version 06.1 2011.

[62] J. Nagumo, S. Arimoto, S. Yoshizawa, *An active pulse transmission line simulating nerve axon.* Proc. IRE Vol. 50, pp. 2061-2070, 1962.

[63] B. Øksendal, *Stochastic differential equations. An introduction with applications.* 6th Edition, Springer-Verlag, Berlin, 2003.

[64] C. Olivera, *Stochastic integration with respect to the cylindrical Wiener process via regularization.* Infin. Dimens. Anal. Quantum Probab. Relat. Top., Vol. 16, (3):9, 2013.

[65] G. A. Pavliotis, A. M. Stuart *Analysis of white noise limits for stochastic systems with two fast relaxation times.* Multiscale Model Simul., Vol. 4, no. 1, pp 1-35, 2005.

[66] A. Pazy, *Semigroups of linear operators and application to partial differential equations.* Applied Mathematical Sciences, Vol. 44, Springer-Verlag, New-York, 1983.

[67] C. Prévôt, M. Röckner, *A concise course on stochastic partial differential equations.* Springer-Verlag, Berlin, 2007.

[68] J. Printems, *On the discretization in time of parabolic stochastic partial differential equations.* Math. Model. and Numer. Anal., Vol. 35, no. 6, pp. 1055-1078, 2001.

[69] M. Pronk, M. Veraar, *A new approach to stochastic evolution equations with adapted drift.* J. Differential Equations 256, no. 11, pp. 3634-3683, 2014.

[70] D. Revuz, M. Yor, *Continuous martingales and Brownian motion.* Springer-Verlag, Berlin, 1999.

[71] J. Rinzel, *Neutrally stable traveling wave solutions of nerve conduction equations.* J. Math. Biol., Vol. 2, no. 3, pp. 205-217, 1975.

[72] J. C. Robinson, *Infinite-dimensional dynamical systems: an introduction to dissipative parabolic PDEs and the theory of global attractors.* Cambridge University Press, 2001.

[73] W. Römisch, R. Winkler *Stochastic DAEs in circuit simulation. Modeling, simulation and optimization of integrated circuits.* Internat. Ser. Numer. Math., 146, pp. 303-318, Birkhäuer, Basel, 2003.

[74] W. Rudin, *Real and complex analysis.* Third Edition, McGraw-Hill Book Co., New York, 1987.

[75] W. Rümelin, *Treatment of stochastic differential equations.* SIAM, Vol. 19, no. 3, pp. 604-613, 1982.

[76] S. Sabanis, *A note on tamed Euler approximations.* Electron. Commun. Probab., Vol. 18, no. 47, 10 pp, 2013.

[77] H. Schurz, *Stochastic α-calculus, a fundamental theorem and Burkholder-Davis-Gundy-type estimates.* Dynam. Syst. Applic., Vol. 15, no. 2, pp. 241-268, 2006.

[78] R. L. Stratonovich *A new representation for stochastic integrals and equations.* SIAM Journal on Control, 4(2), pp. 362-371, 1966.

[79] E. Süli, *Numerical Solution of ordinary differential equations.* Lecture notes for the Mathematical Institute University of Oxford, October 2014.

[80] J. M. Swart, *Pathwise uniqueness for a SDE with non-Lipschitz coefficients.* Stochastic Processes and their Applications, Vol. 98, no. 1, pp.131-149, 2002.

[81] V. Thomée, *Galerkin finite element methods for parabolic problems.* Springer Series in Computational Mathematics, 1997.

[82] R. Tribe, N. Woodward, *Stochastic order methods applied to stochastic travelling waves.* Electron. J. Probab., Vol. 16, no. 16, pp. 436-469, 2011.

[83] J. B. Walsh, *An introduction to stochastic partial differential equations.* École d'été de probabilités de Saint-Flour, XIV-1984, 265-439, Lecture Notes in Math. 1180, Springer, Berlin, 1986.

[84] R. Winkler, *Stochastic differential algebraic equations of index 1 and applications in circuit simulation.* J. Comput. Appl. Math., Vol. 157, no. 2, pp. 477-505, 2003.

[85] K. Yosida, *Functional analysis.* Sixth edition, Springer-Verlag, Berlin- New York, 1980.

[86] Radek Erban, Jonathan Chapman, Ioannis Kevrekidis, Tomas Vejchodsky, *Analysis of a stochastic chemical system close to a SNIPER bifurcation of its mean-field model.* SIAM, Vol. 70, no. 3, pp. 984-1016, 2009.